

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	3474	100.0	868	1	ENV_HV1C4	P05879 human
2	2893.5	83.0	847	1	ENV_HV1S1	P19550 human
3	2869.5	82.6	843	1	ENV_HV1Y2	P35961 human
4	2858	82.3	856	1	ENV_HV1W1	P31872 human
5	2850	82.0	852	1	ENV_HV1S3	P19549 human
6	2836.5	81.6	855	1	ENV_HV1OY	P20888 human
7	2826	81.3	848	1	ENV_HV1JY	P20871 human
8	2819.5	81.2	852	1	ENV_HV1BN	P12488 human
9	2818	81.1	856	1	ENV_HV1SC	P05878 human
10	2815.5	81.0	855	1	ENV_HV1A2	P03378 human
11	2809.5	80.9	856	1	ENV_HV1H2	P04578 human
12	2808	80.8	861	1	ENV_HV1HK	P31819 human
13	2806.5	80.8	856	1	ENV_HV1B1	P03375 human
14	2801.5	80.6	867	1	ENV_HV1J3	P12489 human
15	2799.5	80.6	856	1	ENV_HV1PV	P03376 human
16	2795	80.5	861	1	ENV_HV1BR	P03377 human
17	2788.5	80.3	847	1	ENV_HV1W2	P05880 human
18	2786.5	80.2	865	1	ENV_HV1RH	P04579 human
19	2776.5	79.9	856	1	ENV_HV1H3	P04624 human
20	2773.5	79.8	853	1	ENV_HV1MF	P19551 human
21	2771.5	79.8	856	1	ENV_HV1LW	Q70626 human
22	2768	79.7	851	1	ENV_HV1B8	P04582 human
23	2753.5	79.3	856	1	ENV_HV1MN	P05877 human
24	2623.5	75.5	853	1	ENV_HV1Z2	P12487 human
25	2616.5	75.3	855	1	ENV_HV1Z6	P04580 human
26	2609.5	75.1	863	1	ENV_HV1Z8	P05882 human
27	2577.5	74.2	853	1	ENV_HV1EL	P04581 human
28	2566	73.9	846	1	ENV_HV1ED	P18799 human
29	2550	73.4	856	1	ENV_HV1ZH	P05881 human
30	2516	72.4	859	1	ENV_HV1MA	P04583 human
31	2130.5	61.3	854	1	ENV_STVCZ	P17281 chimera
32	1677.5	48.3	421	1	ENV_HV1N5	P12490 human
33	1616.5	46.5	460	1	ENV_HV1Z3	P12491 human


```
FT CARBOHYD 803 803 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 82.6%; Score 2869.5; DB 1; Length 843;
Best Local Similarity 83.5%; Pred. No. 1.4e-219;
Matches 541; Conservative 35; Mismatches 47; Indels 25; Gaps 6;

Qy 3 LWVTYVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPNPQEVLENVTENF 62
Dy 33 LWVTYVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPNPQEVLENVTENF 92
Qy 63 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTSLTSIIIVVWEQ 122
Dy 93 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTSS---WET 146

Qy 123 RGKEMRNCSENIITTSIRDKVQREYALFYKLDVDFIDDKNNTNTKYRLINCNTSVITQ 182
Dy 147 MEKGEIKNCSENIITTSIRDKVQREYALFYKLDVDFIDDKNNTNTKYRLINCNTSVITQ 199
Qy 183 ACQKVSPEPIPIHYCTPTGFAALLKCNKKFNGTGCNTVSTVQCTHGIRPVVSTOLLING 242
Dy 200 ACQKVSPEPIPIHYCAPAGFAILLKCNKKFNGTGCNTVSTVQCTHGIRPVVSTOLLING 259
Qy 243 SLAEEVIVIRSENFNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVYVYTTGETLG 302
Dy 260 SLAEEVIVIRSENFNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVYVYTTGETLG 319
Qy 303 NROAHNCISRAQNNLTQIATTLRQFG-NKTIAFNQSGGDPPEIWMHSCNGGGEFFY 361
Dy 320 DIRQAHNCISRAQNNLTQIATTLRQFG-NKTIAFNQSGGDPPEIWMHSCNGGGEFFY 379
Qy 362 CNSQLENSANNVTSNGTWSVTRKQKOTGDIITPCRIKQIINWQVGVKAWYALPKGL 421
Dy 380 CNSQLENSANNVTSNGTWSVTRKQKOTGDIITPCRIKQIINWQVGVKAWYALPKGL 429
Qy 422 IRCSNITGLLLTRDGGGENTTIFRPGGDMRDNWSELYKYKVKIEPLGVAPTAK 481
Dy 430 IRCSNITGLLLTRDGGGENTTIFRPGGDMRDNWSELYKYKVKIEPLGVAPTAK 489
Qy 482 RRVQREKRAVGMGAMFLGFLGAAGSTMATVQARQLLSGIVQOQNLLRAIKA 541
Dy 490 RRVQREKRAVGMGAMFLGFLGAAGSTMATVQARQLLSGIVQOQNLLRAIKA 548
Qy 542 QOHLQTLTVWGIKQOARILAVERYLKDQQLLGPWCGSKLICITVAPWNASWKNLTQ 601
Dy 549 QOHLQTLTVWGIKQOARILAVERYLKDQQLLGPWCGSKLICITVAPWNASWKNLTQ 608
Qy 602 IWNNTWMEWDREIDNTYTHLYTTLIESQOQEKNOQELLQDLKQWASL 649
Dy 609 IWNNTWMEWDREIDNTYTHLYTTLIESQOQEKNOQELLQDLKQWASL 656

RESULT 4
ENV HV1W1 STANDARD; PRT; 856 AA.
ID ENV HV1W1
AC P31872;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=31678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=2423250;
RA Starch B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
RT AIDS.";
```

```
Cell 45:637-648 (1986).
-!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.
PIR; A24774; VCLJ3W.
PDB; 1LB0; 04-DEC-02.
PDB; 1LCX; 04-DEC-02.
InterPro; IPR000328; Env GP41.
InterPro; IPR000777; GP120.
Pfam; PF00516; GP120; 1.
Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
3D-structure.
SIGNAL 1 29
CHAIN 30 510
EXTERIOR MEMBRANE GLYCOPROTEIN.
CHAIN 511 856
TRANSMEMBRANE GLYCOPROTEIN.
DIULFID 53 73
BY SIMILARITY.
DIULFID 118 205
BY SIMILARITY.
DIULFID 125 196
BY SIMILARITY.
DIULFID 130 152
BY SIMILARITY.
DIULFID 218 247
BY SIMILARITY.
DIULFID 228 239
BY SIMILARITY.
DIULFID 296 330
BY SIMILARITY.
DIULFID 376 444
BY SIMILARITY.
DIULFID 383 417
BY SIMILARITY.
CARBOHYD 87 87
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 134 134
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 140 140
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 151 151
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 155 155
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 183 183
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 197 197
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 234 234
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 241 241
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 262 262
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 276 276
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 289 289
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 295 295
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 331 331
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 338 338
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 354 354
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 360 360
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 390 390
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 394 394
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 404 404
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 447 447
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 459 459
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 611 611
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 616 616
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 625 625
N-LINKED (GLCNAC...) (POTENTIAL).
CARBOHYD 637 637
N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 856 AA; 97526 MW; DB68D1E49C404DE9 CRC64;

Query Match 82.3%; Score 2858; DB 1; Length 856;
Best Local Similarity 82.8%; Pred. No. 1.2e-218;
Matches 539; Conservative 34; Mismatches 60; Indels 18; Gaps 5;

Qy 3 LWVTYVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPNPQEVLENVTENF 62
Dy 33 LWVTYVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPNPQEVLENVTENF 92
Qy 63 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTSLTSIIIVVWEQ 122
Dy 93 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTSS---ITDEN 140
Qy 123 R---GKEMRNCSENIITTSIRDKVQREYALFYKLDVDFIDDKNNTNTKYRLINCNTSV 179
Dy 141 KTIIGGEVNCSENIITTSIRDKVQREYALFYKLDVDFIDDKNNTNTKYRLINCNTSV 200
Qy 180 ITQACPKVSPEPIPIHYCTPTGFAALLKCNKKFNGTGCNTVSTVQCTHGIRPVVSTOLL 239
Dy 201 ITQACPKVSPEPIPIHYCAPAGFAILLKCNKKFNGTGCNTVSTVQCTHGIRPVVSTOLL 260
```



```

QY 240 LNSGLAEVEVIRSENFNTNAKTIIVQLNVSVVEINCTRPNNHTRKR-VILGPGRVVYTTG 298
Db 261 LNSGLAEVEVIRSENFNTNAKTIIVHLESVVEINCTRPNNVRRHHIHPGRIFY-TG 319
QY 299 EILGNTROAHCHNTRSAQWNTLOQIATTLREQFGNKTIAFNQSSGGDPPEIVMHSFNCGGE 358
Db 320 EIRGNTROAHCHNTRSAQWNTLOQIATTLREQFGNKTIVFNHSSGGDPPEIVHSHFNCGGE 379
QY 359 FFYCNSTQLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAMYALPI 418
Db 380 FFYCDSTQLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAMYALPI 438
QY 419 KGLIRCSSNITGLLLTRDGGENQOTTEIFRPGGDMRDNRWRSYLYKYVKKIBPLGVAPT 478
Db 439 GGQIRCSSNITGLLLTRDGGENQOTTEIFRPGGDMRDNRWRSYLYKYVKKIBPLGVAPT 498
QY 479 KAKRRVVQREKRAVGMGLFGLAAGTGMATSMALTVQARQLLSGIVQOQNNLLRA 538
Db 499 KAKRRVVQREKRAVGAIGAMFLGAGTGMGAASUTLTVQARQLLSGIVQOQNNLLRA 558
QY 539 IKAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWGCGLICTTAVPWNASNKT 598
Db 559 IEAQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWGCGLICTTAVPWNASNKS 618
QY 599 LDOIWNMTWMEWDREIDNTHLYLIESQOQEKNOQELLQDKWASL 649
Db 619 MDQIWNMTWMEWDREIDNTHLYLIESQOQEKNOQELLQDKWASL 669

RESULT 5
ENV_HV1S3
ID ENV HV1S3 STANDARD; PRT; 852 AA.
AC P19549;
RX 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
RT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M38427; AAA45067.1; -.
DR PDB; 1MQQ; 11-DEC-02.
DR HIV; M38427; ENVSSF33.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 507 852 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.

```

```

FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 156 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 331 BY SIMILARITY.
FT DISULFID 377 439 BY SIMILARITY.
FT DISULFID 384 412 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 852 AA; 56663 MW; 5E7BF8D23C9910D CRC64;

```

Query Match 82.0%; Score 2850; DB 1; Length 852;

Best Local Similarity 81.9%; Pred. No. 5,1e-218;

Matches 533; Conservative 51; Mismatches 47; Indels 20; Gaps 7;

```

QY 2 NLWTVYGYGVWVWKEATTLTFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVNTE 61
Db 32 NLWTVYGYGVWVWKEATTLTFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVNTE 91
QY 62 FNNKNNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDL--NTNNTNTTSELSIUV 119
Db 92 FNNKNNMVDQMHEDIISLDQSLKPCVKLTPLCVTLNCTDYLGNATNTNNSGGTV--- 148
QY 120 WEQRGKGMNRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTTNTKYLINCTSV 179
Db 149 ---EKEIKNCSFNITTSIRDKVQREYALFYKLDVVPIDDD-NT--NTSYRLHNCSSV 201
QY 180 ITQACPKVSEPIPHYCTPTGTFALLKNDKXPNGTGCTNVTSTVQCTGIRPVSTOLL 239
Db 202 ITQTCPKVSEPIPHYCTPTGTFALLKNDKXPNGTGCTNVTSTVQCTGIRPVSTOLL 261
QY 240 LNSGLAEVEVIRSENFNTNAKTIIVQLNVSVVEINCTRPNNHTRKVTLPGRVWYTTGE 299
Db 262 LNSGLAEVEVIRSENFNTNAKTIIVQLNVSVVEINCTRPNNHTRKVTLPGRVWYTTGE 321
QY 300 ILGNIRQAHCHNTRSAQWNTLOQIATTLREQFGNKTIAFNQSSGGDPPEIVMHSFNCGGE 359
Db 322 IIGDIRKAYCNISRAKNNKTLQEVATKLEQFGNKTIVFKQSSGGDPPEIVMHSFNCGEF 391
QY 360 FYCNSQTLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAMYALPIK 419
Db 382 FYCNSQTLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAMYALPIK 434
QY 420 GLIRCSSNITGLLLTRDGGENQOTTEIFRPGGDMRDNRWRSYLYKYVKKIBPLGVAPT 478
Db 435 GGQIRCSSNITGLLLTRDGGENQOTTEIFRPGGDMRDNRWRSYLYKYVKKIBPLGVAPT 494

```

QY 479 KAKRRVQREKAVGMLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
 D5 495 KAKRRVQREKAVGMLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 554
 QY 539 IKAQOHLQLTVMGKIQARLAVERYLKQQLLGFWGCCKGLICTTAVPWNASWSNKT 598
 D5 555 IEAQOHLQLTVMGKIQARLAVERYLKQQLLGFWGCCKGLICTTAVPWNASWSNKS 614
 QY 599 LDOIWNNTWWEDEIDNYTHLYTLIESQOQKNEQOELLQDKWASL 649
 D5 615 LDKIWNNTWWEDEIDNYTHLYTLIESQOQKNEQOELLQDKWASL 665

RESULT 6

ENV_HVLOY STANDARD; PRT; 855 AA.

AC P20888;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus
 OX NCBI_TaxID=11699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90148544; PubMed=2559749;
 RA Huot T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
 RT "A highly defective HIV-1 strain isolated from a healthy Gabonese
 RT individual presenting an atypical western blot.";
 RL AIDS 3:707-715(1989).
 CC -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
 CC HEALTHY GABONESE INDIVIDUAL.

CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M26727; AAA83397.1; -
 CC HIV; M26727; ENVSCYI.

DR InterPro; IPR000328; Env GP41.

DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.

DR Pfam; PF00517; GP41; 1.

KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Signal.

FT SIGNAL. 1 29
 FT CHAIN 30 509
 FT CHAIN 510 855
 FT CHAIN 53 73
 FT DISULFID 53 73
 FT DISULFID 118 210
 FT DISULFID 125 201
 FT DISULFID 130 162
 FT DISULFID 223 252
 FT DISULFID 233 244
 FT DISULFID 301 335
 FT DISULFID 381 442
 FT DISULFID 388 415

FT CARBOHYD 87 87
 FT CARBOHYD 134 134
 FT CARBOHYD 142 142
 FT CARBOHYD 145 145
 FT CARBOHYD 161 161
 FT CARBOHYD 165 165
 FT CARBOHYD 192 192
 FT CARBOHYD 202 202

FT EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT TRANSMEMBRANE GLYCOPROTEIN.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.

FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 239 239
 FT CARBOHYD 246 246
 FT CARBOHYD 267 267
 FT CARBOHYD 281 281
 FT CARBOHYD 294 294
 FT CARBOHYD 300 300
 FT CARBOHYD 306 306
 FT CARBOHYD 336 336
 FT CARBOHYD 359 359
 FT CARBOHYD 389 389
 FT CARBOHYD 395 395
 FT CARBOHYD 399 399
 FT CARBOHYD 405 405
 FT CARBOHYD 458 458
 FT CARBOHYD 610 610
 FT CARBOHYD 615 615
 FT CARBOHYD 624 624
 FT CARBOHYD 636 636
 FT CARBOHYD 815 815
 SQ SEQUENCE 855 AA; 97476 MW; 9CF82A607ADD62DA CRC64;

Query Match 81.6%; Score 2836.5; DB 1; Length 855;

Best Local Similarity 81.2%; Pred. No. 6.1e-217;

Matches 532; Conservative 48; Mismatches 50; Indels 25; Gaps 6;

QY 2 NLWTVYGVVPMKEATTTLCASDAKAYDTEAHNVWATHACVPTNPPOEVVLENVTEEN 61

D5 32 NLWTVYGVVPMKEATTTLCASDAKAYATEVNVWATHACVPTNPPOEVVLENVTEEN 91

QY 62 FMWKNMNVQEMHEDIIISLWQSLKPCVKLPLCVTLNCTDINT-----NNTNTTETLSI 116

D5 92 FMWKNMNVQEMHEDIIISLWQSLKPCVKLPLCVTLNCTDINT-----NNTNTTETLSI 150

QY 117 IVWEOQKGEKMRNCSFNITTSIRDKVOREFALFKLDVEPIDDKNTNTNKTRELINCN 176

D5 151 ---WETMEKELKNCSEFNITTSIRDKVOREFALFKLDVEPIDDKNTNTNKTRELINCN 202

QY 177 TSVITQACPKVSFPPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVST 236

D5 203 TSVITQACPKVSFPPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVST 262

QY 237 QLLNGSLAEVEVIRSNFTNNAKTIIVQLNVSVINCTRPNNHNRKRVLTGPGRVMT 296

D5 263 QLLNGSLAEVEVIRSNFTNNAKTIIVQLNVSVINCTRPNNHNRKRVLTGPGRVMT 322

QY 297 TGEILGNIRQAHCLISRAQWNTLQIATTTREQNGKNTIAPNQSOGDPEIWMHSPNCG 356

D5 323 TQKIGDIRQAHCLISRAQWNTLQIATTTREQNGKNTIAPNQSOGDPEIWMHSPNCG 382

QY 357 GEFFVCNSTOLFNSAWNVTSNGTWSVTRKQKQDGTGDIITLPCRIKQIINRWQVVGKAYAL 416

D5 383 GEFFVCNSTOLFNSAWNVTSNGTWSVTRKQKQDGTGDIITLPCRIKQIINRWQVVGKAYAL 434

QY 417 PIKGLIRCSSNITGLLTRDGGGENQTT--BIFRPGGMDRDNWRSELYKYKVKVIEPLG 474

D5 435 PISQIRCSSKITGLLTRDGGGENQTT--BIFRPGGMDRDNWRSELYKYKVKVIEPLG 493

QY 475 VAPTKARRVQREKAVGMLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQN 534

D5 494 VAPTKARRVQREKAVGMLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQN 553

QY 535 LLRAIKAOQHLLQLTVMGKIQARLAVERYLKQQLLGFWGCCKGLICTTAVPWNASW 594

D5 554 LLRAIKAOQHLLQLTVMGKIQARLAVERYLKQQLLGFWGCCKGLICTTAVPWNASW 613

QY 595 SNKTLDOIWNNTWWEDEIDNYTHLYTLIESQOQKNEQOELLQDKWASL 649

D5 614 SNKTLDOIWNNTWWEDEIDNYTHLYTLIESQOQKNEQOELLQDKWASL 668

RESULT 7

ENV_HVLOY

ID - ENV_HVLOY

STANDARD;

PRT; 848 AA.

[illegible]

FT	DISULFID	125	199	BY SIMILARITY.
FT	DISULFID	130	155	BY SIMILARITY.
FT	DISULFID	221	250	BY SIMILARITY.
FT	DISULFID	231	242	BY SIMILARITY.
FT	DISULFID	299	333	BY SIMILARITY.
FT	DISULFID	380	442	BY SIMILARITY.
FT	DISULFID	387	415	BY SIMILARITY.
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	129	129	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	154	154	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	190	190	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	200	200	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	244	244	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	265	265	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	279	279	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	298	298	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	304	304	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	334	334	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	341	341	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	358	358	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	388	388	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	394	394	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	400	400	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	408	408	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	445	445	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	458	458	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	461	461	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	610	610	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	615	615	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	636	636	N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT	CARBOHYD	815	815	N-LINKED (GLCNAC. . .) (POTENTIAL) .
SEQ	SEQUENCE	855 AA;	97438 MW;	A3BC20573AAC41A2 CRC64;

Query Match 81.0%; Score 2815.5; DB 1; Length 855;
 Best Local Similarity 81.2%; Pred. No. 2.8e-215;
 Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6

Qy	3	LWTVTVYGVVPWK	EATTTLFCASDAKAYDTRAHNVWATHACVPTNPQEVLVNTVNF	62
Db	33	LWTVTVYGVVPWK	EATTTLFCASDARAAYDTEVHNWATHACVPTDPNFQEVVLGVNTNF	92
Qy	63	NMWKNNVVEQ	HEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNTNTTNTTSLISVWWEQ	122
Db	93	NMWKNNVVEQ	HEHDIISLWDQSLKPCVKLTPLCVTLNCTDLGKATNTSSN	146
Qy	123	RGKGMENCSPN	ITSIRDKVQREYALFYKLDVBPIDNKNKTTNTTKYRLINCNTSVITQ	182
Db	147	EIKGEIKNCSPN	ITSIRDKIQENALFRNLDVVPIDNASTTNTYTRLIHCNRSVITQ	206
Qy	183	ACPXVSPEPIPIH	YCTPTGFPALLKCNCKPNGTGCTNVSTVQCTHGIRPVVSTQLLNG	242
Db	207	ACPXVSPEPIPIH	YCTPAGFALIKCNKNTFKGPGCTNVSTVQCTHGIRPVVSTQLLNG	266
Qy	243	SLAEEVVIRSENF	TNNAKTTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVYMTTGEILG	302
Db	267	SLAEEVVIRSENF	TNNAKTTIIVQLNESVAINCTRPNNTKRSIYIGPGRAFHTTGRIG	326
Qy	303	NIRQAHCHNISRAQ	NNVILQATTILREOFG-NKTIAFNQSSGGDPEIWMHSPNCGRFFY	361
Db	327	DIRKAHCHNISRAQ	NNVILQEQVKKLREOFGNKNKIVFNQSSGGDPEIWMHSPNCGRFFY	386
Qy	362	CNSTQLFNSA	NWNT-SNGTWSVTRKQDGTIITLPCRIKOLINRWVVGKAMVALPKG	420
Db	387	CNTIQLFNNTRLNH	TEGT-----KGNDT-----IILPCRIKOLINRWVVGKAMVALPKG	438
Qy	421	LIRCSSNITGLLL	TRDGG-GENQTTETPIPRPGGDMRDNWRSLEYKYKVKTEPLGVAPTK	479

FT DISULFID 119 205 BY SIMILARITY.
 FT DISULFID 126 196 BY SIMILARITY.
 FT DISULFID 131 157 BY SIMILARITY.
 FT DISULFID 218 247 BY SIMILARITY.
 FT DISULFID 228 239 BY SIMILARITY.
 FT DISULFID 296 331 BY SIMILARITY.
 FT DISULFID 378 445 BY SIMILARITY.
 FT DISULFID 385 418 BY SIMILARITY.
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match 80.9%; Score 2809.5; DB 1; Length 856;
 Best Local Similarity 82.6%; Pred. No. 8 4e-215;
 Matches 538; Conservative 38; Mismatches 56; Indels 19; Gaps 9;

QY 3 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQVVLNVTFNF 62
 DB 34 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQVVLNVTFNF 93
 QY 63 NMKNMVEQMEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEQ 122
 DB 94 NMKNMVEQMEHDIISLWDQSLKPCVKLTPLCVSLKTDL--KNDTNTSSSGRMIME- 150
 QY 123 RKGEMRNCSEFNITTSIRDKVOREYALFKYLDVEPIDDKNNTNNKXRLNKNVTSVITQ 182
 DB 151 --KGIKNCSEFNISIRKVKQYEAFFKLDIIPDN-----DITSYKLTSCNVSITQ 203
 QY 183 ACPKVSFEPPIHYCTPTGFALLKNDKFNCTGPTNVTSTVQCHGIRPVVSTQLLANG 242
 DB 204 ACPKVSFEPPIHYCAPAGFALLKCNKNTFNCTGPTNVTSTVQCHGIRPVVSTQLLANG 263
 QY 243 SLAESEVIRSENFNTNAKTIIVOLNVSEINCTRNPHTRKRVTL--GPRVWTTTGEI 300
 DB 264 SLAESEVIRSENFNTNAKTIIVOLNVSEINCTRNPHTRKRVTL--GPRVWTTTGEI 323
 QY 301 LGNIQAHNCISRAOWNNTLOQIATTLRBOFG-NKTIAPNQSSGGDPEIVMHSFNCGGEF 359
 DB 324 -GNMQAHNCISRAOWNNTLOQIATTLRBOFG-NKTIAPNQSSGGDPEIVMHSFNCGGEF 382
 QY 360 FYCNSQTQFNSTWNGTSTVTRKQKDTG-DIITLPCRIKQIINRWGVGKAMVALPI 418
 DB 383 FYCNSQTQFNSTW---FNSTWTEGSSNTEGSDTITLPCRIKQIINRWGVGKAMVALPI 439

QY 419 KGLIRCSSNITGLLITRDGGENQTEIFRPGGDMRDNRSELKYKVKIEPLGVAFT 478
 DB 440 SGOIRCSSNITGLLITRDGGENNNESEIFRPGGDMRDNRSELKYKVKIEPLGVAFT 499
 QY 479 KAKRRVVOREKAVGMLGAMFLGFLGAAGSTWGATSMALTVOARQLLSGIVQOONLLRA 538
 DB 500 KAKRRVVOREKAVG-IGALFLGFLGAAGSTWGATSMALTVOARQLLSGIVQOONLLRA 558
 QY 539 IKAQOHLIQLTVWGIKQLOARILAVERYLKDQQLGFWGCSGKLITTAVPNNAWSNKT 598
 DB 559 IEAQOHLIQLTVWGIKQLOARILAVERYLKDQQLGFWGCSGKLITTAVPNNAWSNKS 618
 QY 599 LDOIWNMTMWEDEIDNTHLYLTLEESQOQEKNOQELLQDKWASL 649
 DB 619 LDOIWNMTMWEDEIDNTHLYLTLEESQOQEKNOQELLQDKWASL 669

RESULT 12
 ENV_HV1KB STANDARD; PRT; 861 AA.
 ID ENV_HV1KB
 AC P31819;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 DE GN ENV.
 OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92351552; PubMed=1322587;
 RA Shimizu H., Hasebe F., Tsuchie H., Morikawa S., Ushijima H., Kitamura T.;
 RA "Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated transmembrane glycoprotein.";
 RT Virology 189:534-546(1992).
 RL -!- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN THE CODONS FOR 729-ALA AND 730-ARG.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D12592; BAA02124.1; ALT_SEQ.
 DR PIR; A42995; VCLJKB.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 517 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 518 861 TRANSMEMBRANE GLYCOPROTEIN.
 FT TRANSMEM 690 711 POTENTIAL.
 FT DISULFID 59 79 BY SIMILARITY.
 FT DISULFID 124 212 BY SIMILARITY.
 FT DISULFID 131 203 BY SIMILARITY.
 FT DISULFID 136 160 BY SIMILARITY.
 FT DISULFID 225 254 BY SIMILARITY.
 FT DISULFID 235 246 BY SIMILARITY.
 FT DISULFID 303 337 BY SIMILARITY.
 FT DISULFID 383 446 BY SIMILARITY.
 FT DISULFID 390 419 BY SIMILARITY.
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 643 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 821 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 861 AA; 98116 MW; 306787658F0C9DA CRC64;
Query Match 80.8%; Score 2808; DB 1; Length 861;
Best Local Similarity 81.8%; Pred. No. 1.le-214;
Matches 534; Conservative 49; Mismatches 48; Indels 22; Gaps 9;
QY 3 LWVTYGVGVPMKATTLFCASAKAYDTBAHNVWATHACVPTNPQVVLNVNTNF 62
DB 39 LWVTYGVGVPMKATTLFCASAKAYDTBAHNVWATHACVPTNPQVVLNVNTNF 98
QY 63 NWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCIDLNTNTTTELSIIVWEQ 122
DB 99 NWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCIDLNTNTTTELSIIVWEQ 151
QY 123 RGKEMRNCSENIITSIRKQVREYALFKLDVPEIDDKNTNTNKTLYRLNCTNSVITQ 182
DB 152 KMKGIKNCSENVITIRDKQVREYALFKLDVPEIDDKNTNTNKTLYRLNCTNSVITQ 210
QY 183 ACPKVSFEPIPHYCTPGFALLKNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
DB 211 ACPKVSFEPIPHYCTPGFALLKNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 270
QY 243 SLAEVGVIRSEFTNNAKTIIVQLNVSEINCTPNNHTRKRVTLGPGRVWYTTGRLG 302
DB 271 SLAEVGVIRSEFTNNAKTIIVQLNVSEINCTPNNHTRKRVTLGPGRVWYTTGRLG 330
QY 303 NROAHCHNSRAQNMNTTQQTATTLREOFGNKTIAFNOSGGDPBIVMHSNCGGEFFVC 362
DB 331 DIRQAHCHNSRAQNMNTTQQTATTLREOFGNKTIAFNOSGGDPBIVMHSNCGGEFFVC 390
QY 363 NSTQLFNSAMNVTNSGTWSVTRKQKDTG-DIITLPCRIKQIINRQVVGKAMYPKIGL 421
DB 391 DSSQLFNSTH-LSNGTW-----WNGTGPENTLPCRIKQIINRQVVGKAMYPKIGL 443
QY 422 IRCSSNITGLLITRUGG-----GENQTTIEFRPGGDMRDNRSELYKVKVKEPLGVA 476
DB 444 IRCSSNITGLLITRUGG-----GENQTTIEFRPGGDMRDNRSELYKVKVKEPLGVA 503
QY 477 PTKAKSRVQREKRAVGMGLGAMFLGAGSTMGATSNALTVQARQLLSGTVQOQNLL 536
DB 504 PTKAKSRVQREKRAVGMGLGAMFLGAGSTMGATSNALTVQARQLLSGTVQOQNLL 562
QY 537 RAIKAQHLQLTAVGIIKQLOARILAVERYLKDQQLGFWGSGKLICTTAVPWNASNS 596
DB 563 RAIKAQHLQLTAVGIIKQLOARILAVERYLKDQQLGFWGSGKLICTTAVPWNASNS 622

QY 597 KTLDOINNNMTWMEWDREIDNYTHLYTLTEESNQOEKQKQELLQDLQKWSL 649
DB 623 KSFNEIWDNNMTWMEWDREIDNYTHLYTLTEESNQOEKQKQELLQDLQKWSL 675
RESULT 13
ENV_HV1B1
ID ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumesler K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284 (1985).
[2]
RN DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=9028159; PubMed=2355006;
RX Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
potential glycosylation sites of the type 1 recombinant human
immunodeficiency virus envelope glycoprotein (gp120) expressed in
Chinese hamster ovary cells.";
RL J. Biol. Chem. 265:10373-10382 (1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M15654; AAA44205.1; -.
DR PIR; A03973; VCLJH3.
DR HIV; M15654; ENV\$BHI02.
DR InterPro; IPR000328; ENV_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).

FT CARBOHYD 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 616 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 625 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 637 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 674 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 750 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 816 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 856 AA; 97224 MW; 0BFPIA18931BB27 CRC64;

Query Match 80.8%; Score 2806.5; DB 1; Length 856;
Best Local Similarity 82.8%; Pred. No. 1.5e-214;
Matches 539; Conservative 35; Mismatches 58; Indels 19; Gaps 9;

QY 3 LWVTYGVVPMKEAFTTLFCASDAKAYDEAHNVWATHACVPTNPQEVVLENTENF 62
DB 34 LWVTYGVVPMKEAFTTLFCASDAKAYDEVENWATHACVPTDNPQEVVLENTENF 93
QY 63 NWKKNVQMEHDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEQ 122
DB 94 NWKKNVQMEHDIISLDQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150
QY 123 RKGEMRNCSEFNITTSIRDKQREYALFKLVEPIDDKNTNTNPKYRLNCTSVITQ 182
DB 151 --KGEIKNCSFNISTIRKQVEYAFYKLDIIPDN-----DTTSYLTSCNTSVITQ 203
QY 183 ACPKVSFEPIPIHYCTPTGALLKCNKXFGTGPCTNVSTVQCTHGIRPVVSTQLLLNG 242
DB 204 ACPKVSFEPIPIHYCAPAGFAILLKCNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNG 263
QY 243 SLAEEVVRSENFTNAKTIIVQLNVSEINCTRENNTK--RVTLPGRVWYTTGEI 300
DB 264 SLAEEVVRSENFTNAKTIIVQLNVSEINCTRENNTKRSIRIQRGEGRAVFTIGKI 323
QY 301 LGNIQAHNCISRAQWNTLQOIAATLRBOFG-NKTIAFNQSSGGPPEIWMHSENGGFE 359
DB 324 -GNMRAHNCISRAQWNTLQOIAATLRBOFG-NKTIAFNQSSGGPPEIWMHSENGGFE 382
QY 360 FYCNSQLFNSAWNTVSTGTSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAMVALPI 418
DB 383 FYCNSQLFNSAWNTVSTGTSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAMVALPI 439
QY 419 KGLIRCSSNITGLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVTEPLGVAFT 478
DB 440 KGLIRCSSNITGLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVTEPLGVAFT 499
QY 479 KAKRRVVRKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 500 KAKRRVVRKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 558
QY 539 IKAQHLLOLTVMGIKQLOARILAVERYLKQDQLLFGWCSGKLICTTAVPWNASHSNKT 598
DB 559 IKAQHLLOLTVMGIKQLOARILAVERYLKQDQLLFGWCSGKLICTTAVPWNASHSNKS 618

QY 599 LDOIWNNTWMEWDREIDNYTHLYLTLEESQOQKQOELLQDKWASL 649
DB 619 LDOIWNNTWMEWDREIDNYTHLYLTLEESQOQKQOELLQDKWASL 669

RESULT 14
ENV_HVLJ3 STANDARD; PRT; 867 AA.
ID ENV_HVLJ3
AC P12489;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
GN ENV.
OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11694;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352108; PubMed=2669897;
RA Komiya N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria";
RL AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL; M21138; AAB03526.1; -;
DR HIV; M21138; ENV5JH3.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 217 BY SIMILARITY.
FT DISULFID 125 208 BY SIMILARITY.
FT DISULFID 130 160 BY SIMILARITY.
FT DISULFID 230 259 BY SIMILARITY.
FT DISULFID 240 251 BY SIMILARITY.
FT DISULFID 308 342 BY SIMILARITY.
FT DISULFID 388 457 BY SIMILARITY.
FT DISULFID 395 430 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 867 AA; 98399 MW; 5F2310146B8E8680 CRC64;

Query Match 80.6%; Score 2801.5; DB 1; Length 867;
Best Local Similarity 81.4%; Pred. No. 3.7e-214;
Matches 534; Conservative 41; Mismatches 64; Indels 17; Gaps 7;

Qy 3 LWVTYGVVPVKEATTLFCASDAKAYDTEAHNVWATHACVTPNPNQEVVLENTVF 62
Db 33 LWVTYGVVPVKEATTLFCASDAKAYDTEAHNVWATHACVTPDNPQEVVLENTVEKF 92

Qy 63 NWKNNVVEQHEHDIISLWDSQSLPCVKLTPLCVTLNCTDLNTN---NTNTTSLIIV 119
Db 93 NWKNNVVEQHEHDIISLWDSQSLPCVKLTPLCVTLNCTDLNWDTSNATNTSSG---- 148

Qy 120 WEQRKGEMRNCSENIITTSIDPKVQREYALVKLDVEPID---DNKNTTNTKYLINC 175
Db 149 GEKMEKGMKNCSENIITTSIDPKVQREYALVKLDVEPID---DNKNTTNTKYLISC 208

Qy 176 NTSVITQACPKVSEPIPIHYCTPTGFALLKCNKDKFNGTGPCNTVSTVQCTHGRPVVS 235
Db 209 NTSVITQACPKVSEPIPIHYCAPAGFAIKCNKDKFNGTGPCNTVSTVQCTHGLKPVVS 268

Qy 236 TOLLINGSLABEEVIRSENFTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGGRVWY 295
Db 269 TOLLINGSLABEEVIRSENFTNAKTIIVQLKBPVINCVRPSKTRRRTHIGPGAFY 328

Qy 296 TTGELGNIRAHNCISPAQNNNTLQATTLRQFGNKTIAFNKSSGDPPEIVMHSFNC 355
Db 329 TTKQAGLRQAHNCINRARNATLQVGLKQFVNKTIIVFNKSSGDPPEIVMHSFNC 388

Qy 356 GGEFFYCNSTLFSNMTVSTNGTWSVTRKQDTG--DIITLPCRIKQIINRWQVVGKAM 413
Db 389 GGEFFYCNSTLFSNTW--LSNSTWDTFEGSNNTGGNDITLPCRIKQIINRWQVVGKAM 446

Qy 414 YALPIKGLIRCSNITGLLLTRDGGENQTTETFRPGGDMRDNRWSLYKYVVKVEPL 473
Db 447 YAPPIEGQIRCSNITGLLLTRDGGDNQNETETFRPGGDMRDNRWSLYKYVVKVIELL 506

Qy 474 GVAPTKARRVQREKRAVGMGLGFLGAGTSGATSMALTVOARQLLSGTVOOON 533
Db 507 GVAPTKARRVQREKRAVG--IGAVFLGFLGAGTSGA--SMTLTVQARQLLSGLVQOON 564

Qy 534 NLLRAIKAQHLLQITVWIKQIQLARILAVRYLKQDQLLGFWCSGKLICTTAVPWNAS 593
Db 565 NLLRAIEQOQLLQITVWIKQIQLARILAVRYLKQDQLLGFWCSGKLICTTAVPWNAS 624

Qy 594 WSNKTLQIWNMTWMEWDREIDNYTHLYTLIESQNEQKQOELLQDKWASL 649
Db 625 WSNKSLIEIWNMTWMEWDREIDNYTHLYTLIESQNEQKQOELLQDKWASL 680

RESULT 15
ENV_HV1PV STANDARD; PRT; 856 AA.
AC P03376;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP120 precursor [Contains: Exterior membrane
GN glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
```

```
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCSI_taxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RL AIDS/lymphadenopathy retrovirus.";
RC Nature 313:450-458(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02083; AAB59873.1; -.
DR EMBL; X01762; CAA25903.1; ALT_SEQ.
DR PIR; A03974; VCLJVL.
DR HIV; K02083; ENVSPV22.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97339 MW; 5FCD81DC3C1209B3 CRC64;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:07:21 ; Search time 21 Seconds
(without alignments)
2972.070 Million cell updates/sec

Title: US-09-938-406-1_COPY_33_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVWKEATTT.....QNQEKNOQELLQLDKNASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3474	100.0	868	1 VCLJH4	env polyprotein -
2	2869.5	82.6	843	1 H44001	env polyprotein pr
3	2858	82.3	856	1 VCLJ3W	env polyprotein pr
4	2849.5	82.0	847	2 T09448	envelope glycoprot
5	2837	81.7	852	2 T12016	envelope glycoprot
6	2836.5	81.6	847	2 S13289	env protein - huma
7	2816.5	81.1	852	1 VCLJBR	env polyprotein -
8	2815.5	81.0	855	1 VCLJA2	env polyprotein pr
9	2809.5	80.9	854	2 S13288	env protein - huma
10	2808	80.8	729	1 VCLJXX	env polyprotein pr
11	2808	80.8	861	1 VCLJKB	env polyprotein pr
12	2806.5	80.8	856	1 VCLJH3	env polyprotein pr
13	2799.5	80.6	856	1 VCLJVL	env polyprotein pr
14	2795	80.5	861	1 VCLJLV	env polyprotein pr
15	2782	80.1	851	2 S33985	env polyprotein -
16	2765.5	79.6	861	1 VCLJSC	env polyprotein pr
17	2746	79.0	859	1 VCLJMN	env polyprotein pr
18	2623.5	75.5	853	2 S54384	envelope polyprote
19	2616.5	75.3	855	1 VCLJZR	env polyprotein pr
20	2566	73.9	846	1 VCLJND	env polyprotein pr
21	2550	73.4	856	1 A44963	env polyprotein pr
22	2516	72.4	859	2 T01672	envelope polyprote
23	2130.5	61.3	854	1 VCLJSI	env polyprotein pr
24	2074	59.7	506	2 A40218	envelop glycoprote
25	1909	55.0	495	2 S31493	env polyprotein -
26	1825	52.5	443	2 C41621	env polyprotein P
27	1802	51.9	445	2 A41621	env polyprotein M
28	1751.5	50.4	877	2 S49197	envelope protein p
29	1746.5	50.3	454	2 B41621	env polyprotein D

ALIGNMENTS

RESULT 1

VCLJH4

env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)

N/Alternate names: coat polyprotein

N/Contains: coat protein gp120; coat protein gp41

C/Species: human immunodeficiency virus type 1, HIV-1

C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999

C/Accession: C25523

R/Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S.

Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986

A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human 3

A/Reference number: A94136; MUID:87041461; PMID:3490666

A/Accession: C25523

A/Molecule type: DNA

A/Residues: 1-868 <DB>

A/Cross-references: GB:M13137; MID:g326460; PIDN:AAA44311.1; PID:g326467

C/Genetics:

A/Gene: env

C/Superfamily: type E retrovirus env polyprotein

C/Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei

F.1-521/Product: coat protein gp120 #status predicted <GP1>

F.522-868/Product: coat protein gp41 #status predicted <GP2>

F.89,138,139,142,166,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,

Query Match 100.0%; Score 3474; DB 1; Length 868;

Best Local Similarity 100.0%; Pred. No. 2e-247;

Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANLWTVYGVVWKEATTTLCASDAKAYDEAHNVWATHACVPTNENPOEVLENVTE 60

DB 33 ANLWTVYGVVWKEATTTLCASDAKAYDEAHNVWATHACVPTNENPOEVLENVTE 92

QY 61 NFNMKNMVMQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVW 120

DB 93 NFNMKNMVMQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVW 152

QY 121 EQRKGEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVI 180

DB 153 EQRKGEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVI 212

QY 181 TQACPVSFEPPIHYCTPTGTFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLL 240

DB 213 TQACPVSFEPPIHYCTPTGTFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLL 272

QY 241 NGSIAEEVWIRSENFNTNAKTIIVQLNVSEINCTRNHTRKRVTLGPGRVATTGEI 300

DB 273 NGSIAEEVWIRSENFNTNAKTIIVQLNVSEINCTRNHTRKRVTLGPGRVATTGEI 332

QY 301 LGNIRQAHCNISRAQWNNTLQOIATTLREQFGNKTIAPNQSSGGDPPEIVMHSFNCGGEFF 360

DB 333 LGNIRQAHCNISRAQWNNTLQOIATTLREQFGNKTIAPNQSSGGDPPEIVMHSFNCGGEFF 392

QY 361 YCNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKG 420
 DB 393 YCNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKG 452
 QY 421 LIRCSSNITGALLTRDGGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAK 480
 DB 453 LIRCSSNITGALLTRDGGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAK 512
 QY 481 KRRVQREKRAVGMGLGAMFGLGAGSTMGATSMALTQAROLLGSGIVQOQNLLRAIK 540
 DB 513 KRRVQREKRAVGMGLGAMFGLGAGSTMGATSMALTQAROLLGSGIVQOQNLLRAIK 572
 QY 541 AQOHLQLTVMGIKOLQARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLTD 600
 DB 573 AQOHLQLTVMGIKOLQARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLTD 632
 QY 601 QIWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 649
 DB 633 QIWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 681

RESULT 2
 H44001
 env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
 N;Alternate names: coat polyprotein
 N;Contains: coat protein gp120; coat protein gp41
 C;Species: human immunodeficiency virus type 1, HIV-1
 A;Note: host Homo sapiens (man)
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
 C;Accession: H44001
 J;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
 J. Virol. 66, 6587-6600, 1992
 A;Title: Complete nucleotide sequence, genome organization, and biological properties of
 A;Reference number: A44001; MUID:93021387; PMID:1404605
 A;Accession: H44001
 A;Molecule type: DNA
 A;Residues: 1-843 <LTY>
 A;Cross-references: GB:M93258
 C;Genetics:
 A;Gene: env
 C;Superfamily: type E retrovirus env polyprotein
 C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F;1-29/Domain: signal sequence #status predicted <SIG>
 F;19-35/Region: hydrophobic
 F;30-48/Product: coat protein gp120 #status predicted <GP1>
 F;490-843/Product: coat protein gp41 #status predicted <GP2>
 F;499-515/Region: hydrophobic
 F;673-689/Region: hydrophobic
 F;738-755/Domain: transmembrane #status predicted <TMN>
 F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 82.6%; Score 2869.5; DB 1; Length 843;
 Best Local Similarity 83.5%; Pred. No. 5.7e-203;
 Matches 541; Conservative 35; Mismatches 47; Indels 25; Gaps 6;

QY 3 LWTVVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
 DB 33 LWTVVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
 QY 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNTNTTTELSIIVVWEQ 122
 DB 93 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDL--RNAINTSSS---WET 146
 QY 123 RGKEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTKYRLINCNTSVITQ 182
 DB 147 MEKGEIKNCSPNITTSIRDKVQREYALFYKLDVEPID--NASTYRLINCNTSVITQ 199
 QY 183 ACQKVSFEPIPIHYCTPTGFALLKCNCKNGTGPCINVTQCTHGIRPVVSTQILLG 242
 DB 200 ACQKVSFEPIPIHYCAPAGFALLKCNCKNGTGPCINVTQCTHGIRPVVSTQILLG 259
 QY 243 SLAEEFVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGELG 302

DB 260 SLAEEFVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKSINIGPRALYTTGELIG 319
 QY 303 NIRAHCNISRANWNTLQOIATLREOFG-NKTIARNQSSGGDPEIVMHSFNCGGFFY 361
 DB 320 DIRQAHCNLSKQWENTLEQIAIKLKEQFQNNKTIIFNPSSGGDPEIVTSHFNCGGFFY 379
 QY 362 CNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKGL 421
 DB 380 CNSTOLF-----TWNDTRKLNNTGNITLPCRIKOIINRWQVGVKAMVALPIRQG 429
 QY 422 IRCSSNITGALLTRDGGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAK 481
 DB 430 IRCSSNITGALLTRDGGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAK 489
 QY 482 RRVVQREKRAVGMGLGAMFGLGAGSTMGATSMALTQAROLLGSGIVQOQNLLRAIKA 541
 DB 490 RRVVQREKRAVG-LGALFLGFLGAGSTMGAAITLTVQAROLLGSGIVQOQNLLRAIEA 548
 QY 542 QOHLQLTVMGIKOLQARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLTD 601
 DB 549 QOHLQLTVMGIKOLQARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKLINE 608
 QY 602 IWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 649
 DB 609 IWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 656

RESULT 3
 VCLJ3W
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)
 N;Contains: coat protein gp120; coat protein gp41
 C;Species: human immunodeficiency virus type 1, HIV-1
 C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 07-Nov-1997
 C;Accession: A24774
 R;Stargich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.S.
 Cell 45, 637-648, 1986
 A;Title: Identification and characterization of conserved and variable regions in the en
 A;Reference number: A24774; MUID:86218077; PMID:2423250
 A;Accession: A24774
 A;Molecule type: DNA
 A;Residues: 1-856 <STA>
 A;Cross-references: GB:K03455; GB:M38432; NID:gl906382
 C;Genetics:
 A;Gene: env
 C;Superfamily: type E retrovirus env polyprotein
 C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
 F;1-29/Domain: signal sequence #status predicted <Sig>
 F;30-501/Product: coat protein gp120 #status predicted <GP1>
 F;502-847/Product: coat protein gp41 #status predicted <GP2>
 F;87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459,

Query Match 82.3%; Score 2858; DB 1; Length 856;
 Best Local Similarity 82.8%; Pred. No. 4.1e-202;
 Matches 539; Conservative 34; Mismatches 60; Indels 18; Gaps 5;

QY 3 LWTVVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
 DB 33 LWTVVYGVVPWKATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
 QY 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNTNTTTELSIIVVWEQ 122
 DB 93 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLN-----ITDWN 140
 QY 123 R---GKEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTKYRLINCNTSV 179
 DB 141 KTIIGGEVKNCSFNITTSIRDKVQREYALFYKLDVVPPIKSNNDSTYTRLRHCHNTSV 200
 QY 180 ITOACKVSFEPIPIHYCTPTGFALLKCNCKNGTGPCINVTQCTHGIRPVVSTQLL 239
 DB 201 ITOACKVSFEPIPIHYCAPAGFALLKCNCKNGTGPCINVTQCTHGIRPVVSTQLL 260
 QY 240 LNSLAEEFVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTG 298

Db 261 LNSLAEEBIVIRSENFTDNAKTIIIVHLSVEINCTRNNNVRRBHHIHPGRAPY-TG 319
QY 299 EILGNTRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGE 358
Db 320 EIRGNTRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGE 379
QY 359 FFYCNSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPI 418
Db 380 FFYCNSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPI 438
QY 419 KGLIRCSSNITGLLLTRDGGNGQTTTFRPGGDMRDNRWSLYKYKVVKIEPLGVAPT 478
Db 439 GGOIRCSSNITGLLLTRDGGNGQTTTFRPGGDMRDNRWSLYKYKVVKIEPLGVAPT 498
QY 479 KAKRRVVQREKRAVGMFLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRA 538
Db 499 KAKRRVVQREKRAVGMFLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRA 558
QY 539 IKAQCHLLQLTVMGIKQIOLARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWNSKT 598
Db 559 IEAQCHLLQLTVMGIKQIOLARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWNSKT 618
QY 599 LDQIWNMTWMEWRIDNVTSLIYTLIBESQOQEKNOQELLDKQWASL 649
Db 619 MDQIWNMTWMEWRIDNVTSLIYTLIBESQOQEKNOQELLDKQWASL 669

RESULT 4
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09448
Rifang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A:Reference number: Z16673
A:Accession: T09448
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-847 <PAN>
C:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 82.0%; Score 2849.5; DB 2; Length 847;
Best Local Similarity 82.8%; Pred. No. 1.7e-201;
Matches 536; Conservative 40; Mismatches 52; Indels 19; Gaps 4;
QY 3 LMTVYVYGVPMWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62
Db 33 LMTVYVYGVPMWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 92
QY 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWBEQ 122
Db 93 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWBEQ 146
QY 123 RQKGMNCSFNITTSIRDKVQREYALFYKLDVPEIDDKNKTNTNKTLYELNCTSVITQ 182
Db 147 -ERGBIKNSFNITTSIRDEQVEYALFYKLDVPEIDDKNKTNTNKTLYELNCTSVITQ 200
QY 183 ACQKVSFEPIHYCTPTGFCALLKNDKFKNGTGPCTNVSTVQCTHGRPVVSTQLLNG 242
Db 201 ACQKVSFEPIHYCAPAGFALLKNDKFKNGTGPCTNVSTVQCTHGRPVVSTQLLNG 260
QY 243 SLAEEVIRSENFTNNAKTIIVQLNVSVEINCTRNHTRKRVTLGPRVWYTTG 302
Db 261 SLAEEVIRSENFTNNAKTIIVQLNVSVEINCTRNHTRKRVTLGPRVWYTTG 320
QY 303 NTRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGEFFYC 362
Db 321 DIRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGEFFYC 380

QY 363 NSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 422
Db 381 NSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 434
QY 423 RCSSNITGLLLTRDGGNGQTTTFRPGGDMRDNRWSLYKYKVVKIEPLGVAPTAKR 482
Db 435 RCSSNITGLLLTRDGGNGQTTTFRPGGDMRDNRWSLYKYKVVKIEPLGVAPTAKR 494
QY 483 RVQREKRAVGMFLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRAIKAQ 542
Db 495 RVQREKRAVGMFLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRAIKAQ 553
QY 543 QHLLQLTVMGIKQIOLARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWNSKTLDQI 602
Db 554 QHLLQLTVMGIKQIOLARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWNSKTLDQI 613
QY 603 WNNMTWMEWRIDNVTSLIYTLIBESQOQEKNOQELLDKQWASL 649
Db 614 WNNMTWMEWRIDNVTSLIYTLIBESQOQEKNOQELLDKQWASL 660

RESULT 5
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T12016
R.McCutchan, F.B.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S
A:Reference number: Z17379; MUID:98178716; PMID:9519894
A:Accession: T12016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-852 <MCC>
C:Cross-references: EMBL:U09934; NID:g2351783; PID:NAC59271.1; PID:g2351784
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.7%; Score 2837; DB 2; Length 852;
Best Local Similarity 83.1%; Pred. No. 1.4e-200;
Matches 541; Conservative 35; Mismatches 53; Indels 22; Gaps 8;
QY 3 LMTVYVYGVPMWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62
Db 33 LMTVYVYGVPMWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 92
QY 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWBEQ 122
Db 93 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWBEQ 143
QY 123 RG- - -KGMNCSFNITTSIRDKVQREYALFYKLDVPEIDDKNKTNTNKTLYELNCTNS 178
Db 144 NKGSGEGEIKNCSFNITTRIGNKVKQYALFYKLDVPEIDDKNKTNTNKTLYELNCTNS 198
QY 179 VITQACPKVSFEPIHYCTPTGFCALLKNDKFKNGTGPCTNVSTVQCTHGRPVVSTQ 238
Db 199 VITQACPKVSFEPIHYCTPTGFCALLKNDKFKNGTGPCTNVSTVQCTHGRPVVSTQ 258
QY 239 LINGSLAEEVIRSENFTNNAKTIIVQLNVSVEINCTRNHTRKRVTLGPRVWYTTG 298
Db 259 LINGSLAEEVIRSENFTNNAKTIIVQLNVSVEINCTRNHTRKRVTLGPRVWYTTG 318
QY 299 ETLGNIRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGE 358
Db 319 ALIGDIRQAHCHNISRAQWNTTQQIATTIREPGNKTIAFNQSSGGDPEIVMHSFNCGE 378
QY 359 FFYCNSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPI 418
Db 379 FFYCNSTQLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVGVKAMVALPI 435
QY 419 KGLIRCSSNITGLLLTRDGGNGQTTTFRPGGDMRDNRWSLYKYKVVKIEPLGVAPT 478

Db 436 RGQIRCSNITGILLTRDGGTNNSTNETFRPGGDMRDNWRSELYKYVKVIEPLGVAPT 495
Qy 479 KAKRVTVQREKRAVGMFLGALGAGSTMGATSMALTVOARQLLSGIVQQNNLLRA 538
Db 496 KAKRVTVQREKRAVG-IGALFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNNLLRA 554
Qy 539 IKAQOHLQLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWKNKT 598
Db 555 IEAQOHLQLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWKNKS 614
Qy 599 LDQTNWNTWMEWDREIDNYTHLYTLEESQOQKNOQELLQDLKWSL 649
Db 615 LDKIWNNTWMEWDREIDNYTHLYTLEESQOQKNOQELLQDLKWSL 665

RESULT 6
S13289
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13289
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13289
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <ORF>
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.6%; Score 2836.5; DB 2; Length 847;
Best Local Similarity 82.5%; Pred. No. 1.5e-200;
Matches 534; Conservative 39; Mismatches 55; Indels 19; Gaps 4;

Qy 3 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 33 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
Qy 63 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDNTNTTNTTNTLSIIVVWEQ 122
Db 93 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDVNATNTDSEGTM----- 146

Qy 123 RGKEMRNCSEFNITTSIRDKVQREYALFYKLDVBPIDNKNNTNTKYLINCMSTVITQ 182
Db 147 -ERGEIKNCSEFNITTSIRDEYQKEALFYKLDVBPIDN-----NNTSYRLISCDSVTITQ 200

Qy 183 ACPKVSFEPIPIHYCTPTGFPALLKNDKKFGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 201 ACPKISFEPIPIHYCAPAGFAILLKNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLNG 260

Qy 243 SLAEEVVRSENFNTNAKTIIVQLNVSEINCRPNNHTKRVTLPGRVYMTTGEILG 302
Db 261 SLAEEVVRSENFNTNAKTIIVQLKESVEINCRPNNHTKRVTLPGRVYMTTGEILG 320

Qy 303 NIRAHCNISRAQNNLTQIATTLREQFGKNTAFNOSGGDPEIIVMHSFCGGEFFYC 362
Db 321 DIRAHCNISRAKNDTLKQIVILREQFENKTIIVFNHSSGGDPEIIVMHSFCGGEFFYC 380

Qy 363 NSTQLFSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAPIKGLI 422
Db 381 NSTQLNNTSWNNTGEGS-----NNTGNTITLPCRIKQIINRWQVGVKAMVAPIRQI 434

Qy 423 RCSSNITGLLTRDGGGNNOTTEIFRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKR 482
Db 435 RCSSNITGLLTRDGGINGENTEIFRPGGDMRDNWRSELYKYVKVIEPLGVAPTAKR 494

Qy 483 RVQREKRAVGMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAIKAO 542
Db 495 RVQREKRAVG-IGAVFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNNLLRAIEAQ 553

Qy 543 QHLLQLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWKNKTLDOI 602

Db 554 QRLQLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWKNKSLDRI 613
Qy 603 WNNMTWMEWDREIDNYTHLYTLEESQOQKNOQELLQDLKWSL 649
Db 614 WNNMTWMEWDREIDNYTHLYTLEESQOQKNOQELLQDLKWSL 660

RESULT 7
VCLJBR
env polyprotein - human immunodeficiency virus type 1 (isolate BR)
N:Alternate names: coat polyprotein
N:Contains: coat protein gp120; coat protein gp41
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
C:Accession: A31667
R:Ahand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.
Virology 168, 79-89, 1989
A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)
A:Reference number: A94389; MUID:89085613; PMID:2789516
A:Accession: A31667
A:Molecule type: DNA
A:Residues: 1-852 <ANA>
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: capsid protein; coat protein; polyprotein; transmembrane protein
F:1-516/Product: coat protein gp120 #status predicted <CP>
F:517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 81.1%; Score 2816.5; DB 1; Length 852;
Best Local Similarity 80.6%; Pred. No. 4.6e-199;
Matches 528; Conservative 53; Mismatches 43; Indels 31; Gaps 7;

Qy 3 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 34 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 93

Qy 63 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDNTNTTNTTNTLSIIVVWEQ 122
Db 94 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCHDFNATNTS----- 143

Qy 123 RGK-----GEMRNCSEFNITTSIRDKVQREYALFYKLDVBPIDNKNNTNTKYLINCMST 178
Db 144 -GKWMGEGEMKNCSEFNITTSIRDRMQEYALFYKLDVBPIDNK--TNTYRLISCNTS 199

Qy 179 VITQACPKVSFEPIPIHYCTPTGFPALLKNDKKFGTGPCTNVSTVQCTHGIRPVVSTQL 238
Db 200 VITQACPKVSFEPIPIHYCAPAGFAILLKNDKKFGTGPCTNVSTVQCTHGIRPVVSTQL 259

Qy 239 LLNGLAEEVVRSENFNTNAKTIIVQLNVSEINCRPNNHTKRVTLPGRVYMTTGEILG 298
Db 260 LLNGLAEEVVRSENFNTNAKTIIVQLNESVEINCRPNNHTKRVTLPGRVYMTTGEILG 319

Qy 299 EILGNIRQAHCNISRAQNNLTQIATTLREQFGKNTAFNOSGGDPEIIVMHSFCGGE 358
Db 320 QIIGDIRRAHCNLSRKWENTLKQIVTKLRVQPKNTIVFNHSSGGDPEIIVMHSFCGGE 379

Qy 359 PFYCNSTQLFNSAW-NVTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAP 416
Db 380 PFCNTTQLFNSTVYRNTTGNIT-----EGNSPTLPCRIKQIINRWQVGVKAMVAP 431

Qy 417 PIKGLIRCSNITGLLTRDGGGNNOT--TEIFRPGGDMRDNWRSELYKYVKVIEPLG 474
Db 432 PIRQIKCSSNITGLLTRDGGNNNTTDEIFRPGGDMRDNWRSELYKYVKVIEPLG 491

Qy 475 VAPTAKRVRVQREKRAVGMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNN 534
Db 492 VAPTAKRVRVQREKRAVG-LGALFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNN 550

Qy 535 LLRAIKAOQHLQLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASW 594
Db 551 LLMAIEAQOQHLMLTVMGIKOLQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASW 610

QY 595 SNKTLQIWNMTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649
Db 611 SNKSLSDIWNMTWMEWDREIDNYTHLYIYTLIEESQIOEKNEKELLELDKWASL 665

RESULT 8
VCLJA2
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03976
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Sh
Science 227, 484-492, 1985
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A:Reference number: A04003; MUID:85090453; PMID:2578227
A:Accession: A03976
A:Molecule type: DNA
A:Residues: 1-855 <SAN>
A:Cross-references: GB:K02007; NID:G328658; PIDN:AAB59882.1; PID:G328666
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro
F:1-30/Domain: signal sequence #status predicted <Sig>
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>
F:87,129,140,158,184,190,200,244,265,292,298,304,334,341,358,364,388,394,400,408,445,458
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.0%; Score 2815.5; DB 1; Length 855;
Best Local Similarity 81.2%; Pred. No. 5.4e-199;
Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;

QY 3 LWTVYGVVPVWKEATTTLCFASDADKAYDTEAHNVWATHACVPTNPNQEVVLNVTFNF 62
Db 33 LWTVYGVVPVWKEATTTLCFASDADKAYDTEAHNVWATHACVPTDNPQEVVLNVTFNF 92

QY 63 NWKKNMVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTNTELSIIVWFQ 122
Db 93 NWKKNMVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146

QY 123 RKGEMRNCSPNITTSIRDKVOREVALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 182
Db 147 EIKGEIKNCSPNITTSIRDKIQENALFNLVDVPIDNASTTTNTYNYRLIHCRNSVITQ 206

QY 183 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 207 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPIVSTQLLNG 266

QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSVINCTRNHNTRKVTILGPRVWYTTGEILG 302
Db 267 SLAEVWIRSENFNTNAKTIIVQLNVSVINCTRNHNTRKSIYIGPGRAFTTGRIIG 326

QY 303 NTRQAHNCISRAQWNTLQOIATTLRBOFG-NKTIAFNQSSGGDPPIVHVSFNCGGEPPY 361
Db 327 DIRXAHNCISRAQWNTLQOIATTLRBOFG-NKTIAFNQSSGGDPPIVHVSFNCGGEPPY 386

QY 362 CNSTOLFNSAWNT-SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAWYALPIKG 420
Db 387 CNTQTFNNTWRINHTEGT-----KNDT---IILPCRIKQIINRWQVGVKAWYALPIGG 438

QY 421 LIRCSSNITGLLITRDGG-GENOTTEIFRPGGDMRDNRSSELYKYKVKIPELGVAPT 479
Db 439 QISCSSNITGLLITRDGGINVTNDTEVFRPGGDMRDNRSSELYKYKVIKIPGLIAPT 498

QY 480 AKRRVVQREKRAVGMIGAMFLGFGAAGSTMGTATSMALTVOARQLLSGIVQOQNLLRAI 539
Db 499 AKRRVVQREKRAVGMIGAMFLGFGAAGSTMGTATSMALTVOARQLLSGIVQOQNLLRAI 558

QY 540 KAOQHLLQLTVMGIKQOARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNSKTL 599
Db 540 KAOQHLLQLTVMGIKQOARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNSKTL 599

Db 559 EAQOHLQLTVMGIKQOARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNSKSL 618
QY 600 DOIWNMTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649
Db 619 EDIWNMTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 668

RESULT 9
S13288
env protein - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S13288
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Digne, A.; Idler, K.; Zack, J.A.
Nature 348, 69-73, 1990
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c
A:Reference number: S13288; MUID:91043044; PMID:2172833
A:Accession: S13288
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-854 <OBR>
C:Superfamily: type E retrovirus env polyprotein

Query Match 80.9%; Score 2809.5; DB 2; Length 854;
Best Local Similarity 82.8%; Pred. No. 1.5e-198;
Matches 539; Conservative 34; Mismatches 57; Indels 21; Gaps 9;

QY 3 LWTVYGVVPVWKEATTTLCFASDADKAYDTEAHNVWATHACVPTNPNQEVVLNVTFNF 62
Db 34 LWTVYGVVPVWKEATTTLCFASDADKAYDTEAHNVWATHACVPTDNPQEVVLNVTFNF 93

QY 63 NWKKNMVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTNTELSIIVWFQ 122
Db 94 NWKKNMVQEMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDTNTNSSSGRMIME- 150

QY 123 RKGEMRNCSPNITTSIRDKVOREVALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 182
Db 151 --KGIEIKNCSPNITTSIRDKVOREVALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 201

QY 183 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 202 ACPKVSFEPIPIHYCAPAGFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 261

QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSVINCTRNHNTRK--RVTLGPGRVWYTTGEI 300
Db 262 SLAEVWIRSENFNTNAKTIIVQLNVSVINCTRNHNTRKSIYIGPGRAFTTGRI 321

QY 301 LGNIRQAHNCISRAQWNTLQOIATTLRBOFG-NKTIAFNQSSGGDPPIVHVSFNCGGEFF 359
Db 322 -GNMIRQAHNCISRAQWNTLQOIATTLRBOFG-NKTIAFNQSSGGDPPIVHVSFNCGGEFF 380

QY 360 FYCNSDTOLFNSAWNT-SNGTWSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAWYALPI 418
Db 381 FYCNSDTOLFNSAWNT-SNGTWSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAWYALPI 437

QY 419 KGLIRCSSNITGLLITRDGGINVTNDTEVFRPGGDMRDNRSSELYKYKVKIPELGVAPT 478
Db 438 SQGIRCSSNITGLLITRDGGINVTNDTEVFRPGGDMRDNRSSELYKYKVKIPELGVAPT 497

QY 479 KAKRRVVQREKRAVGMIGAMFLGFGAAGSTMGTATSMALTVOARQLLSGIVQOQNLLRA 538
Db 498 KAKRRVVQREKRAVGMIGAMFLGFGAAGSTMGTATSMALTVOARQLLSGIVQOQNLLRA 556

QY 539 IKAQOHLQLTVMGIKQOARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNSKTL 598
Db 557 IKAQOHLQLTVMGIKQOARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNSKTL 616

QY 599 LDOIWNMTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649
Db 617 LDOIWNMTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 667

RESULT 10

VCLJHX

env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp32
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996
C;Accession: B42995
R;Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.
Virology 189, 534-546, 1992
A;Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A;Reference number: A42995; MUID:92351552; PMID:1322587
A;Accession: B42995
A;Molecule type: mRNA
A;Residues: 1-729 <SH1>
A;Cross-references: GB:S41266; GB:D01206
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-689/Domain: extracellular #status predicted <EX1>
F;1-33/Domain: signal sequence #status predicted <SIG>
F;17-33/Region: hydrophobic #status predicted
F;34-517/Product: coat protein gp120 #status predicted <CP1>
F;514-517/Region: cleavage processing #status predicted
F;518-729/Product: coat protein gp32 #status predicted <CP2>
F;518-729/Region: hydrophobic #status predicted
F;538-534/Region: transmembrane #status predicted <TM1>
F;690-711/Domain: transmembrane #status predicted <INT>
F;712-729/Domain: intracellular #status predicted <INT>
F;93, 141, 146, 146, 163, 191, 192, 237, 241, 248, 269, 283, 296, 308, 338, 345, 361, 367, 397, 403, 408, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

Query Match	80.8%;	Score 2808;	DB 1;	Length 729;
Best Local Similarity	81.8%;	Pred. No. 1.6e-198;		
Matches 534;	Conservative 49;	Mismatches 48;	Indels 22;	Gaps 9
Qy	3	LWTVVYGV	PWKATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTNF	62
Db	39	LWTVVYGV	PWKATITLFCASDAKAYDTEAHNVWATHACVPTDPNQEVLVNVTNF	98
Qy	63	NMWKNNV	VEQMHEDIISLWDQSLKPCVKLPLCVLTNCIDLNTNNTTETLSIIVWBEQ	122
Db	99	NMWKNNV	VEQMHENIIISLWDQSLKPCVKLPLCVTLHCTDL--RNTTNNS--SI---EE	155
Qy	123	RGKGMENC	SEFNITTSIRDVKQREYALPKYLDVEPIDDKNTNNTKVRLLNCNTSVITQ	182
Db	152	XMKGEI	KNCSENVITNRDKVKQYALPKYLDVPIDNDNST--NTCVRLISCDSIVITQ	210
Qy	183	ACPVSFE	PIPIHYCTPTGTFALLKNDCKFNGTGPGCTNVSTVQCTHGIRPVVSTOLLNG	242
Db	211	ACPVSFE	PIPIHYCTPTAGFALLKCNKNTFNGTGPCKNVSTVQCTHGIRPVVSTQLLNG	270
Qy	243	SLABEEV	VRSENTTNAKTIIVOLNVSEVNCITRPNNHTRKRVTLGPRVWYTTGEILG	302
Db	271	SLABEGV	VRSENTFDNVKTIIVQLNETVKINCIRPNKTRKRVTMGPRVYTTGEILG	330
Qy	303	NIRQAH	CNISRAQNNTLQOIATTLREQFGNKTAFNOSGGDPEIVMHSEFCGGEFFYC	362
Db	331	DIRQAH	CNISRAEWNKLTLEQIANLKRQFENKTIIVFNOSGGDPEIVMHSEFCGGEFFYC	390
Qy	363	NSITOLF	SNAMVNTSGVTSVTRKODTG-DIITLPCRIOIINRWQVVGKAWYALPIKGL	421
Db	391	DSQLFN	STH--LNGTW-----WNGTGPENITLPCRIOIINMWQEVGKAWYAPPIRQ	443
Qy	422	IRGSSNT	IGLLTRDGG-----GNQTTPEIFRPGGDMRDNRWSRLYKVKVVKIEPLGVA	476
Db	444	IRCSSN	ITGLLTRDGGNTNNNTNSSITEIFRPGGDMRDNRWSRLYKVKVVKIEPLGVA	503
Qy	477	PTKAKRV	QREKEAVGMLGAMFLGFLGAAGSTWGAATVQARQLISGIVQOONLL	536
Db	504	PTKAKRV	QREKEAVG--IGAVFLGFLGAAGSTWGAATVTVQARQLLPGIVQOONLL	562
Qy	537	RAIKAO	QHLLQLTVMGIKQLOARIIAVERYIKDQQLLGFWCSCGKLICTTAPVPMNASWN	596

Dbb

Qy

Dbb

RESULT 11
VCLJKB
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gg)
N;Alternate names: coat polypotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 01-Mar-1996
R/Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura,
Viology 189, 534-546, 1992
A;Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a
A;Reference number: A42995; MUID:92351552; PMID:1322587
A;Accession: A42995
A;Molecule type: mRNA
A;Residues: 1-861 <SHI>
A;Cross-references: GB:S41266; GB:D01206

C/Genetics: 1
A/Gene: env
C/Superfamily: type E retrovirus env polypeptide
C/Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F/1-689/Domain: extracellular #status predicted <EXT>
F/1-33/Domain: signal sequence #status predicted <SIG>
F/17-33/Region: hydrophobic #status predicted
F/34-517/Product: coat protein gp120 #status predicted <CPI>
F/514-517/Region: cleavage processing #status predicted
F/518-861/Product: coat protein gp41 #status predicted <CP2>
F/518-861/Region: hydrophobic #status predicted
F/690-711/Domain: transmembrane #status predicted <TM1>
F/712-861/Domain: intracellular #status predicted <INT>
F/756-772/Region: hydrophobic #status predicted
F/93_141_145_146_163_191_192_237_241_248_269_283_296_308_338_345_361_367_397_414

[illegible]

Db 444 IRCSSNITGLLLTRDGGNTQNNNTSSIEIFRPGGDMRDNRSELYKYKVKIEPLGVA 503
Qy 477 PTKAKRRVQREKRAVGLGAMFLGFGAAGSTMGATSMALTIVQARQLLSGIVQQQNLL 536
Db 504 PTKAKRRVQREKRAVG-IGAVFLGFLGAAGSTMGAAVTLTVQARQLLPQIVQQQNLL 562
Qy 537 RAIKAQHLQLTVMGIKOLQARILAVERYLKDOQLLGFWCSGKLICTTAVPNNASHN 596
Db 563 RAIDAOQHLQLTVMGIKOLQARILAVERYLKDOQLLGFWCSGKLICTTAVPNNASHN 622
Qy 597 KTLQIWNMTWMEWDREIDNTHLYIETIESNQKNOQELLQDLKWSL 649
Db 623 KSFNEIWNMTWMEWEREINNTLYNIETIESNQKNOQELLQDLKWSL 675

RESULT 12
VCLJH3
env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03973
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora-
nberger, J.A.; Papas, T.S.; Graybe, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
A:Reference number: A93353; MUID:85111123; PMID:2578615
A:Accession: A03973
A:Molecule type: DNA
A:Residues: 1-856 <RAT>
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro-
tein; 1-30/Domain: signal sequence #status predicted <SIG>
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 80.8%; Score 2806.5; DB 1; Length 856;
Best Local Similarity 82.8%; Pred. No. 2.5e-198;
Matches 539; Conservative 35; Mismatches 58; Indels 19; Gaps 9;

Qy 3 LWVTYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62
Db 34 LWVTYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 93
Qy 63 NWNKNNVQMHEDIISLWDSQKPCVKLTPLCVLTNCTDLNNTNTTSLSIIVVWEQ 122
Db 94 NWNKNDVQMHEDIISLWDSQKPCVKLTPLCVLSKCTDL--KNDTNTSSSGRMIME- 150
Qy 123 RGKGMNCSFNITSIIRDKVQREYALFYKLDVPEIDNKNNTNTKYLINCNSTVITQ 182
Db 151 --KGEIKNCSFNITSIIRGKVKQKEYAFYKLDIIPDN----DTTSYLTSCNTSVITQ 203
Qy 183 ACPKVSFPIPIHYCTPTGFALLKCNKNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 204 ACPKVSFPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 263
Qy 243 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVVYTTGEI 300
Db 264 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVVYTTGEI 323
Qy 301 LGNTRQAHCNISRAQWNNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHSENCGGEF 359
Db 324 -GNRQAHCNISRAQWNNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHSENCGGEF 382
Qy 360 FYCNSTQLFNSW---FNSTWSTKGSNNTSGSDITLPCRILQIINMMQEVGKAWYALPI 418

Db 383 FYCNSTQLFNSW---FNSTWSTKGSNNTSGSDITLPCRILQIINMMQEVGKAWYALPI 439
Qy 419 KGLIRCSSNITGLLLTRDGGGENTTEIFRPGGDMRDNRSELYKYKVKIEPLGVAPT 478
Db 440 SGQIRCSSNITGLLLTRDGGGNSNNEBIFRPGGDMRDNRSELYKYKVKIEPLGVAPT 499
Qy 479 KAKRRVQREKRAVGLGAMFLGFGAAGSTMGATSMALTIVQARQLLSGIVQQQNLLRA 538
Db 500 KAKRRVQREKRAVG-IGAVFLGFLGAAGSTMGAAVTLTVQARQLLSGIVQQQNLLRA 558
Qy 539 IKAQHLQLTVMGIKOLQARILAVERYLKDOQLLGFWCSGKLICTTAVPNNASHNKT 598
Db 559 IEAQHLQLTVMGIKOLQARILAVERYLKDOQLLGFWCSGKLICTTAVPNNASHNKS 618
Qy 599 LDQIWNMTWMEWDREIDNTHLYIETIESNQKNOQELLQDLKWSL 649
Db 619 LEQIWNMTWMEWDREINNTLYNIETIESNQKNOQELLQDLKWSL 669

RESULT 13
VCLJVL
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C:Accession: A03974
R:Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.
Nature 313, 450-458, 1985
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi-
A:Reference number: A93355; MUID:85111157; PMID:2982104
A:Accession: A03974
A:Molecule type: DNA
A:Residues: 1-856 <MOE>
A:Cross-references: GB:K02083; NID:g555008; PIDN:AA859873.1; PID:g328559
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprote-
in; 1-30/Domain: signal sequence #status predicted <SIG>
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 80.6%; Score 2799.5; DB 1; Length 856;
Best Local Similarity 82.5%; Pred. No. 8.2e-198;
Matches 537; Conservative 36; Mismatches 59; Indels 19; Gaps 9;

Qy 3 LWVTYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62
Db 34 LWVTYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 93
Qy 63 NWNKNNVQMHEDIISLWDSQKPCVKLTPLCVLTNCTDLNNTNTTSLSIIVVWEQ 122
Db 94 NWNKNDVQMHEDIISLWDSQKPCVKLTPLCVLSKCTDL--KNDTNTSSSGRMIME- 150
Qy 123 RGKGMNCSFNITSIIRDKVQREYALFYKLDVPEIDNKNNTNTKYLINCNSTVITQ 182
Db 151 --KGEIKNCSFNITSIIRGKVKQKEYAFYKLDIIPDN----DTTSYLTSCNTSVITQ 203
Qy 183 ACPKVSFPIPIHYCTPTGFALLKCNKNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 204 ACPKVSFPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 263
Qy 243 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVVYTTGEI 300
Db 264 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVVYTTGEI 323
Qy 301 LGNTRQAHCNISRAQWNNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHSENCGGEF 359
Db 324 -GNRQAHCNISRAQWNNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHSENCGGEF 382
Qy 360 FYCNSTQLFNSW---FNSTWSTKGSNNTSGSDITLPCRILQIINMMQEVGKAWYALPI 418

QY 360 FYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMVALPI 418
Db 383 FYCNSTOLFNSMTW---FNSTWSTEGSNNTGSDTITLPCRIKQIINRWQVGVKAMVAPPI 439
QY 419 KGLIRCSSNITGLLLTRDGGGNETTEIFRPGGDMRDNWRSELYKYVVKVIEPLGVAPT 478
Db 440 SQGIQCSSNITGLLLTRDGGGNNNESEIFRPGGDMRDNWRSELYKYVVKVIEPLGVAPT 499
QY 479 KAKRRVVQREKRAVGMFLGFLGAAGTGMATSMALTQVQRLLSGIVQOQNLLRA 538
Db 500 KAKRRVVQREKRAVG-IGALFLGFLGAAGTGMATSMALTQVQRLLSGIVQOQNLLRA 558
QY 539 IKAQOHLQLTWVGILQVQRIILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASNKT 598
Db 559 IEAQOHLQLTWVGILQVQRIILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASNKS 618
QY 599 LDOIWNMTWMEWDREIDNTHLYTLIERSONQOEKNOQELQLDKWASL 649
Db 619 LEQIWNMTWMEWDREIDNTHLYTLIERSONQOEKNEQELLELDKWANL 669
RESULT 14
VCLJLV
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C;Accession: A03975
R;Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.
Cell 40, 9-17, 1985
A;Title: Nucleotide sequence of the AIDS virus, LAV.
A;Reference number: A90866; MUID:85099333; PMID:2981635
A;Accession: A03975
A;Molecule type: DNA
A;Residues: 1-861 <WAI>
A;Cross-references: GB:K02013; NID:g326417; PIDN:AA859751.1; PID:g326424
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-516/Product: exterior membrane glycoprotein #status predicted <EXT>
F;517-861/Product: transmembrane glycoprotein #status predicted <TM>
F;88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411
F;616,621,630,642,679,755,821/Binding site: carboxydrate (Asn) (covalent) #status predic
Query Match 80.5%; Score 2795; DB 1; Length 861;
Best Local Similarity 82.4%; Pred. No. 1.8e-197;
Matches 539; Conservative 37; Mismatches 58; Indels 20; Gaps 10;
QY 3 LWTVYGVVPVWEKATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 62
Db 34 LWTVYGVVPVWEKATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 93
QY 63 NMKNMNVQEMHEDIISLWDQSLKPCVKLPCLVTLNCTDIN--TN--NTNTTSLTIIV 119
Db 94 NMKNMNVQEMHEDIISLWDQSLKPCVKLPCLVTLNCTDIN--TN--NTNTTSLTIIV 153
QY 120 WEQKGMENCSFNITTSIRDKVQREYALFYKLDVEFDIDNKNNTNNTKYRLINCNVT 179
Db 154 ME---KGEIKNCNSFNISTSRGKQVQKEYAFYKLDIIPDN----DTTSYLTSCNTSV 205
QY 180 ITQACPKVSEPIPIHYCTGTGALLKCNKDKFNGTGPCTNVSTVQCTHGIRPWSVQTL 239
Db 206 ITQACPKVSEPIPIHYCAPAGEFALLKCNKDKFNGTGPCTNVSTVQCTHGIRPWSVQTL 265
QY 240 LNSLAEEVVRSENFNTNNAKTIIVQLNVSVSEINCTRPNNHTK--RVTLGPRVWYTT 297
Db 266 LNSLAEEVVRSENFNTNNAKTIIVQLNVSVSEINCTRPNNHTK--RVTLGPRVWYTT 325
QY 298 GEILGNIRQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNQSOGGDEPEIVHSEFCG 356

Db 326 GKI-GNMROAHNCISRAQWNNTLQIATTLREQFGNKNKTIIFKQSSGGDEPEIVHSEFCG 384
QY 357 GEFFYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMYA 415
Db 385 GEFFYCNSTOLFNSMTW---FNSTWSTEGSNNTGSDTITLPCRIKQIINRWQVGVKAMYA 441
QY 416 LPTKGLIRCSSNITGLLLTRDGGGNETTEIFRPGGDMRDNWRSELYKYVVKVIEPLGV 475
Db 442 PPSGQIRCSSNITGLLLTRDGGGNNNGSEIFRPGGDMRDNWRSELYKYVVKVIEPLGV 501
QY 476 APTKAKRRVVQREKRAVGMFLGFLGAAGTGMATSMALTQVQRLLSGIVQOQNLL 535
Db 502 APTKAKRRVVQREKRAVG-IGALFLGFLGAAGTGMATSMALTQVQRLLSGIVQOQNLL 560
QY 536 LRAIKAQOHLQLTWVGILQVQRIILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASWS 595
Db 561 LRAIEAQOHLQLTWVGILQVQRIILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASWS 620
QY 596 NKTLDQIWNMTWMEWDREIDNTHLYTLIERSONQOEKNOQELQLDKWASL 649
Db 621 NKSLEQIWNMTWMEWDREIDNTHLYTLIERSONQOEKNEQELLELDKWASL 674
RESULT 15
S33985
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C;Accession: S33985
R;Carlini, F.
submitted to the EMBL Data Library, November 1991
A;Reference number: S33979
A;Accession: S33985
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-851 <CAR>
A;Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199
C;Superfamily: type E retrovirus env polyprotein
Query Match 80.1%; Score 2782; DB 2; Length 851;
Best Local Similarity 82.0%; Pred. No. 1.6e-196;
Matches 534; Conservative 39; Mismatches 54; Indels 24; Gaps 9;
QY 3 LWTVYGVVPVWEKATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 62
Db 34 LWTVYGVVPVWEKATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 93
QY 63 NMKNMNVQEMHEDIISLWDQSLKPCVKLPCLVTLNCTDIN--TN--NTNTTSLTIIV 122
Db 94 NMKNMNVQEMHEDIISLWDQSLKPCVKLPCLVTLNCTDIN--TN--NTNTTSLTIIV 150
QY 123 RGKGMENCSFNITTSIRDKVQREYALFYKLDVEFDIDNKNNTNNTKYRLINCNVT 182
Db 151 --KGEIKNCNSFNISTSRGKQVQKEYAFYKLDIIPDN----DTTSYLTSCNTSV 203
QY 183 ACQKVPSEPIPIHYCTGTGALLKCNKDKFNGTGPCTNVSTVQCTHGIRPWSVQTL 242
Db 204 ACQKVPSEPIPIHYCAPAGEFALLKCNKDKFNGTGPCTNVSTVQCTHGIRPWSVQTL 263
QY 243 SLAEEVVRSENFNTNNAKTIIVQLNVSVSEINCTRPNNHTK--RVTLGPRVWYTT 300
Db 264 SLAEEVVRSENFNTNNAKTIIVQLNVSVSEINCTRPNNHTK--RVTLGPRVWYTT 323
QY 301 LGNIRQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNQSOGGDEPEIVHSEFCG 359
Db 324 -GNMROAHNCISRAQWNNTLQIATTLREQFGNKNKTIIFKQSSGGDEPEIVHSEFCG 382
QY 360 FYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMVALPI 418
Db 383 FYCNSTOLFNSMTW---FNSTWSTEGSNNTGSDTITLPCRIKQIINRWQVGVKAMVAPPI 434
QY 419 KGLIRCSSNITGLLLTRDGGGNETTEIFRPGGDMRDNWRSELYKYVVKVIEPLGVAPT 478

Db 435 SGOIRGSSNITGLLTRDGGNSNESEIFRPGGDMRDNRSELYKYXVVKIEPLGVAPT 494
QY 479 KAKRRVVQREKRAVGM/GAMFLGFLGAAGSTMGATSMALTVOAROLLSGIVQOQNNLLRA 538
Db 495 KAKRRVVQREKRAVG--IGALFLGFLGAAGSTMGAASMTLTVOAROLLSGIVQOQNNLLRA 553
QY 539 IKAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASWSNKT 598
Db 554 IEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWGCSGKPICTTAVPWNASWSNKS 613
QY 599 LDCIWNNTWMEWDREIDNYTHIYTLIBESQOQEKNOQELLQLDKWASL 649
Db 614 LEQIWNNTWMEWDREINNYTSLIHSLIBESQOQEKNEQELLQLDKWASL 664

Search completed: January 13, 2004, 09:11:58
Job time : 23 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 08:59:06 ; Search time 48 Seconds
(without alignments)
2146.116 Million cell updates/sec

Title: US-09-938-406-1_COPY_33_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVWKEATT.....QNQKEKNQELLQLDKWASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3457	99.5	649	19 AAW44250	HIV-1 gp160 residu
2	3388	97.5	865	19 AAW43072	HIV-1 gp120 protei
3	2977.5	85.7	883	22 AAB82761	Ancestral HIV-1 gr
4	2883.5	83.0	847	21 AAY97073	Variant HIV-1 SF16
5	2878.5	82.9	842	24 ABU66565	Human immunodefici
6	2870.5	82.6	842	23 ABB06211	HIV Env isolate SF
7	2852	82.1	850	16 AAR67724	gp120 from the HIV
8	2834	81.6	855	18 AAW11581	Human Immunodefici
9	2834	81.6	855	20 AAW88113	Env protein of the

10	2827.5	81.4	643	22 AAB61505	HIV-1 SOS gp140 gl
11	2816	81.1	851	9 AAP80367	HIV protein Hf6.
12	2815.5	81.0	855	19 AAW53112	ENV protein contai
13	2815.5	81.0	855	21 AAY77298	HIV-1 (ATCC CRL 85
14	2815.5	81.0	855	21 AAY77302	HIV-1 (ATCC CRL 85
15	2815.5	81.0	855	24 ABU57550	AIDS associated re
16	2815.5	81.0	855	24 ABU57553	AIDS associated re
17	2815.5	81.0	863	7 AAP61509	Sequence of ARV-2
18	2815.5	81.0	863	13 AAP29706	env gene decoded f
19	2814.5	81.0	854	21 AAB10697	HIV-1 env protein.
20	2814.5	81.0	854	21 AAB10053	HIV-1 coat protei
21	2814.5	81.0	854	22 AAB86199	HIV gp41 DNA SEQ I
22	2814.5	81.0	854	23 ABG73663	HIV-1 NL4-3 gp120
23	2814	81.0	880	18 AAW23333	Human immunodefici
24	2814	81.0	880	20 AAW73332	HIV envelope prote
25	2814	81.0	880	21 AAB14844	HIV envelope prote
26	2812	80.9	857	12 AAR12261	HIV-1 strain OVI E
27	2811.5	80.9	854	22 AAB67277	Protein encoded by
28	2809.5	80.9	856	21 AAY97072	Wild type HIV-1 HX
29	2809.5	80.9	857	16 AAR67725	gp120 from the HIV
30	2806.5	80.8	856	22 AAB85999	Amino acid sequenc
31	2806.5	80.8	863	14 AAR43869	HTLV-III ENV-LOR g
32	2804	80.7	848	23 AAU11872	HIV env protein fr
33	2801	80.6	860	18 AAW31284	HIV-SF2 virus gp12
34	2799.5	80.6	856	7 AAP61514	Sequence of envelo
35	2799.5	80.6	856	20 AAW89325	HIV-1 env protein
36	2798.5	80.6	856	13 AAR25940	Modified HIV env g
37	2796.5	80.5	863	7 AAP60349	HTLV-III virus (H
38	2795.5	80.5	856	22 AAB45697	HIV-1/IIIB env clo
39	2795	80.5	868	23 AAO19389	Lymphadenopathy-as
40	2792	80.4	868	7 AAP60422	Sequence of LAV vi
41	2791	80.3	751	10 AAP90604	HIV-1 env segment
42	2791	80.3	856	16 AAR67726	gp120 from the HIV
43	2790.5	80.3	856	7 AAP60131	Sequence of the AI
44	2788.5	80.3	866	9 AAP80966	HIV protein HT7.
45	2788	80.3	868	7 AAP60063	HIV virus env gene

ALIGNMENTS

RESULT 1

AAW44250
ID AAW44250 standard; protein; 649 AA.

XX AAW44250;

XX 26-JUN-1998 (first entry)

XX HIV-1 gp160 residues 33-681.

DE Vaccine; antibody; antigen; hydrophobic; proteosome; pathogen;
KW immune response; sexually transmitted disease; HIV; infection.

OS Human immunodeficiency virus type 1.

XX Key Location/Qualifiers

FT Region 491..519
/note= "hydrophobic region"

FN WO9801558-A2.

XX 15-JAN-1998.

XX 10-JUL-1997; 97WO-US12253.

XX 10-JUL-1996; 96US-0021687.

XX (INTE-) INTELLIVAX INC.

PA (JACK-) JACKSON FOUND HENRY M.

XX (USSA) US SEC OF ARMY.

PI Birx DL, Lowell GH, Vancott TC;

Db 32 ANLWTVYGVPKKEA-TTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTE 90
 QY 61 NFNMMKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNNTN-TTSLIIV 119
 Db 91 NFNMMKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNNTN-TTSLIIV 150
 QY 120 WEORGKGMRCNSFNITTSIRDKVOREYALFYKLDVEPIDDDKN-TTNNTKYLINCN 178
 Db 151 WEORGKGMRCNSFNITTSIRDKVOREYALFYKLDVEPIDDDKN-TTNNTKYLINCN 210
 QY 179 VITQCPKVSFPEPIHYCTPTGFALLKCNCKKFGTGTCTNVSTVQCHGIRPVVSTOL 238
 Db 211 VITQCPKVSFPEPIHYCTPTGFALLKCNCKKFGTGTCTNVSTVQCHGIRPVVSTOL 270
 QY 239 LLNGSLAEVEVIRSENFNNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPRVWYTG 298
 Db 271 LLNGSLAEVEVIRSENFNNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPRVWYTG 330
 QY 299 EILGNIRQAHONISRAQWNNTLQOATTLREOFGNKTIAFNOSGGDPEIVMHSFNC 358
 Db 331 QILGNIRQAHONISRAQWNNTLQOATTLREOFGNKTIAFNOSGGDPEIVMHSFNC 390
 QY 359 FFCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQIINRWGVGKAMVALPI 418
 Db 391 -FFCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQIINRWGVGKAMVALPI 448
 QY 419 KGLIRCSSNITGLLTRDGGGNETTEIPRPGGDMRDNRSELYKYKVKVTEPLGAPT 478
 Db 449 KGLIRCSSNITGLLTRDGGGNETTEIPRPGGDMRDNRSELYKYKVKVTEPLGAPT 508
 QY 479 KAKRRVVOREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQQNNLLRA 538
 Db 509 KAKRRVVOREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQQNNLLRA 567
 QY 539 IKAQOHLQLTVMGKIQOLQARILAVERYLKDQQLLGFWSGSKLICTTAVPWNASWSNKT 598
 Db 568 IKAQOHLQLTVMGKIQOLQARILAVERYLKDQQLLGFWSGSKLICTTAVPWNASWSNKT 627
 QY 599 LQIWNNTWMEWDREIDNYTHLIYTLIEESQKQKQQLLQDKWASL 649
 Db 628 LQIWNNTWMEWDREIDNYTHLIYTLIEESQKQKQQLLQDKWASL 678

RESULT 3

AAB82761
 ID AAB82761 standard; Protein; 883 AA.

AC AAB82761;
 DT 29-OCT-2001 (first entry)

XX Ancestral HIV-1 group M, subtype B gp160 protein.
 DE HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.

XX Human immunodeficiency virus type 1.
 OS WO200160838-A2.
 FN 23-AUG-2001.

PD 16-FEB-2001; 2001WO-US05288.
 PF 18-FEB-2000; 2000US-0183659.

XX (UNIW) UNIV WASHINGTON.
 PA Mullins JI, Rodrigo AG, Learn GH, Li F;

PI WPI; 2001-536565/59.
 DR N-PSDB; AAB26468.

XX

PT Preparing an ancestral viral amino acid sequence useful as a vaccine
 PT comprises determining a recent ancestor of a circulating virus by
 XX maximum likelihood phylogeny analysis -
 PS Claim 8; Page 54; 89pp; English.

CC The present sequence is that of an ancestral HIV-1 group M, subtype
 CC B gp160 (env gene product) sequence. The invention provides
 CC compositions and methods for determining ancestral viral gene
 CC sequences and ancestral viral protein sequences for highly diverse
 CC viruses, such as HIV-1. The methods use samples of circulating
 CC viruses to determine an ancestral viral sequence by maximum
 CC likelihood phylogeny analysis. In the present case, the ancestral
 CC HIV-1 subtype B env sequence (see AAB26468) was determined using 38
 CC subtype B sequences (obtained from 9 different countries) and 3
 CC subtype D (outgroup) sequences. The distances between this ancestral
 CC viral sequence and circulating strains used to determine it were on
 CC average 12.3% (range: 8.0-21.0%) while the available specimens were
 CC 17.3% different from each other (range: 13.3-23.2%). Thus, the
 CC ancestor sequence was, on average, more closely related to any given
 CC circulating virus than to any other variant. The ancestral gp160
 CC sequence included a wide variety of immunogenically active peptides
 CC when processed for antigen presentation; nearly all known subtype B
 CC CTL epitope consensus amino acids were represented. Thus, an
 CC immunogenic composition to this subtype B ancestor protein will
 CC elicit broad neutralising antibody against HIV-1 isolates of the
 CC same subtype, and will also elicit a broad cellular response
 CC mediated by antigen-specific T-cells. A claimed vaccine composition
 CC comprises a viral ancestor protein or its immunogenic fragment,
 CC especially one derived from the HIV-1 group M subtype B gp160
 CC ancestral protein.

XX Sequence 883 AA;

Query Match 85.7%; Score 2977.5; DB 22; Length 883;
 Best Local Similarity 84.7%; Pred. No. 5.7e-160;

Matches 564; Conservative 27; Mismatches 54; Indels 21; Gaps 6;

QY 3 LWTVYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTE 62

Db 33 LWTVYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTE 92

QY 63 NFNMMKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNNTTSLIIVWEQ 122

Db 93 NFNMMKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNNTTSLIIVWEQ 152

QY 123 RG-----KGEMRCNSFNITTSIRDKVOREYALFYKLDVEPID-DNKNTNNTKYRLIN 174

Db 153 SGGTMEGEKGEIKNCSENVITTSIRDKVOREYALFYKLDVEPIDDNNTNNTSYRLIN 212

QY 175 CNTSVITQACPKVSFPEPIHYCTPTGFALLKCNCKKFGTGTCTNVSTVQCHGIRPVV 234

Db 213 CNTSVITQACPKVSFPEPIHYCTPTGFALLKCNCKKFGTGTCTNVSTVQCHGIRPVV 272

QY 235 STOLLNGSLAEVEVIRSENFNNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPRVW 294

Db 273 STOLLNGSLAEVEVIRSENFNNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPRVW 332

QY 295 YTTGEILGNIRQAHONISRAQWNNTLQOATTLREOFGNK-TTAFNOSGGDPEIVMHS 352

Db 333 YATGKIIGDIRQAHONISRAQWNNTLQOATTLREOFGNKTTIVFNOSGGDPEIVMHS 392

QY 353 FNCGBEFFYCNSTOLFNSAMNVTSGTWSVRKQK-----DTGDIITLPCRIKQIINRWQ 407

Db 393 FNCGBEFFYCNSTOLFNSAMNVTSGTWSVRKQK-----DTGDIITLPCRIKQIINRWQ 450

QY 408 VVGKAMVALPIKGLIRCSSNITGLLTRDGGGNETTEIPRPGGDMRDNRSELY 463

Db 451 EVGKAMVAPPISGQIRCSNITGLLTRDGGGNETTEIPRPGGDMRDNRSELY 510

QY 464 KYKVVKIEPLGVAPTAKRRVVOREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQ 523

Db 511 KYKVVKIEPLGVAPTAKRRVVOREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQ 570

QY 524 LLSGIVQQNNLLRAIKAAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSCGLI 583
DB 571 LLSGIVQQNNLLRAIEAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSCGLI 630
QY 584 CTTAVPNAWSNKTLDQIWNMTWMEDEIDNYTHLYTLIEESQOQKNEQELQL 643
DB 631 CTTAVPNAWSNKSLLDKIWNMTWMEDEIDNYTHLYTLIEESQOQKNEQELQL 690
QY 644 DKWASL 649
DB 691 DKWASL 696

RESULT 4
AAAY97073
ID AAAY97073 standard; protein; 847 AA.
XX
AC AAAY97073;
XX
DT 31-OCT-2000 (first entry)
XX
DE Variant HIV-1 SF162 Env gp160.
XX
KW HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;
KW CD4 binding region; VI/V2 loop; bridging sheet.
XX
OS Human immunodeficiency virus type 1 isolate SF162.
XX
EN WO200039303-A2.
XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31272.
XX
PR 31-DEC-1998; 98US-0114495.
PR 29-SEP-1999; 99US-0156670.
XX
PA (CHIR) CHIRON CORP.
XX
PI Barnett S, Hartog K, Martin E;
XX
DR WPI; 2000-465745/40.
XX
PT Novel modified HIV Env polypeptides useful as immunizing agents and for
PT preparing a vaccine to elicit an immune response against a broad range
PT of HIV subtypes
XX
PS Claim 5; Page 115-117; 139pp; English.

XX
CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope
CC (Env) polypeptides having an amino acid deleted or replaced in the region
CC corresponding to residues 420-436 or 119-123 and 199-210 relative to
CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV
CC strain SF162, with numbering relative to isolate HXB-2. The Env
CC polypeptides are modified so as to expose at least part of the CD4
CC binding region. The modified HIV Env polypeptides, coding polynucleotides
CC and constructs, further comprising an adjuvant, are used for inducing an
CC immune response in an individual. The method involves administering a
CC first composition comprising a polynucleotide encoding the Env
CC polypeptide in a priming step and administering a second composition
CC comprising a modified Env polypeptide as a booster in an amount
CC sufficient to induce an immune response in the individual. The first
CC and/or second composition further comprises an adjuvant (claimed). The
CC intracellularly produced Env polypeptides can be used for a number of
CC diagnostic and therapeutic purposes to determine the presence of reactive
CC antibodies/and or Env proteins in a biological sample to aid in the
CC diagnosis of HIV infection or disease status or as measure of response to
CC immunization.

XX SQ Sequence 847 AA;

Query Match 83.0%; Score 2883.5; DB 21; Length 847;

Best Local Similarity 83.3%; Pred. No. 1.le-154;
Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;
QY 3 LWTVTVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPDPQEVLVNTENF 62
DB 33 LWTVTVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPDPQEVLVNTENF 92
QY 63 NMKNVWVQMHEDIISLWDSLKPCVKLTPLCVTLNCTDLNNTNTTNTTSLIIVVWEQ 122
DB 93 NMKNVWVQMHEDIISLWDSLKPCVKLTPLCVTLNCTDLNNTNTTSSN----WKE 146
QY 123 RGKGMENCSFNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTNTKYRLINCMTSVITQ 182
DB 147 MDRGEIKNCSEKVTTSIRNKQKEYALFYKLDVVPIDN-----DNTSYKLINCNTSVITQ 201
QY 183 ACPKVSPEPIPIHYCTPTGFAILLKCNCKKPNFGPCPTNVSTVQCTHGIRPVVSTQILLNG 242
DB 202 ACPKVSPEPIPIHYCAPAGFAILLKCNCKKPNFGPCPTNVSTVQCTHGIRPVVSTQILLNG 261
QY 243 SLAEVEVIRSENFNTNAKTIIVQLNVSVENCTRPNNHTRKRVTLGPGRVWYTTGELG 302
DB 262 SLAEVGVIRSENFNTNAKTIIVQLNVSVENCTRPNNHTRKRVTLGPGRVWYTTGELG 321
QY 303 NIROAHNCISRAOWNTLQIATTLREQFGNKTIAFNQSSGGDPEIYVHGFNCGGEFFYC 362
DB 322 DIRQAHNCISGEKWNNTLKQIVTKLQAFGNKTIIVFKQSSGGDPEIYVHGFNCGGEFFYC 381
QY 363 NSTQLFNSANVT-----SNGTWSVTRKQDGTIILPCRIKQIINRWQVVGKAMALP 417
DB 382 NSTQLFNSANVT-----SNGTWSVTRKQDGTIILPCRIKQIINRWQVVGKAMALP 428
QY 418 IKGLIRCSSNITGLLTLRDGGE-NOTTEIFRPGGDMRDNWRSELYKYVVKTEPLGVA 476
DB 429 IRGQIRCSSNITGLLTLRDGGE-NOTTEIFRPGGDMRDNWRSELYKYVVKTEPLGVA 488
QY 477 PTKAKRRVQREKRAVGMGLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLL 536
DB 489 PTKAKRRVQREKRAV-TLGNMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLL 547
QY 537 RAIKAQOHLQLTWGKIQARILAVERYIKDQQLLGFWGCSCGLICTTAVPNAWSN 596
DB 548 RAIKAQOHLQLTWGKIQARILAVERYIKDQQLLGFWGCSCGLICTTAVPNAWSN 607
QY 597 KTDQIWNMTWMEDEIDNYTHLYTLIEESQOQKNEQELQLDKWASL 649
DB 608 KSLDQIWNMTWMEDEIDNYTHLYTLIEESQOQKNEQELQLDKWASL 660

RESULT 5

ABU66565
ID ABU66565 standard; Protein; 842 AA.
XX
AC ABU66565;
XX
DT 22-MAY-2003 (first entry)
XX
DE Human immunodeficiency virus (HIV) envelope (env) protein #1.
XX
KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;
KW vaccine; gene therapy; packaging cell line; humoral immune response;
KW cellular immune response; gene delivery vector; DNA immunisation;
KW envelope protein; env.
XX
OS Human immunodeficiency virus.
XX
PN WO2003004657-A1.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002WO-US21421.
XX
PR 05-JUL-2001; 2001US-303192P.
PR 31-AUG-2001; 2001US-316860P.

PR 16-JAN-2002; 2002US-349728P.
PR 16-JAN-2002; 2002US-349793P.
PR 16-JAN-2002; 2002US-349871P.
XX
XX
XX (CHIR) CHIRON CORP.
XX
XX Zur Megede J, Barnett SW, Lian Y;
XX WPI; 2003-221602/21.
XX
XX
XX New synthetic polynucleotides encoding antigenic HIV type B and/or type
XX C polypeptides, useful as immunogenic compositions or vaccines for
XX generating humoral or cellular immune responses against HIV in a
XX subject, especially humans -
XX
XX Disclosure; Fig 2A-C; 262pp; English.
XX
XX The invention describes a synthetic polynucleotide encoding 2 or more
XX immunogenic HIV polypeptides, where at least 2 of the polypeptides are
XX derived from different HIV subtypes. The polynucleotide is useful for
XX immunisation, generation of packaging cell lines, or production of HIV
XX polypeptides. The polynucleotide and its encoded proteins are useful as
XX immunogenic compositions or vaccines for generating humoral or cellular
XX immune responses against HIV in a subject, or for inducing neutralising
XX antibodies against HIV. The gene delivery vector comprising the
XX polynucleotide is also useful for DNA immunisation of, or for
XX generating an immune response (e.g. a humoral or cellular immune
XX response) in, a subject such as a mammal, particularly a human. This
XX is the amino acid sequence of a human immunodeficiency virus (HIV)
XX envelope (env) protein.
XX
XX Sequence 842 AA;
XX
Query Match 82.9%; Score 2878.5; DB 24; Length 842;
Best Local Similarity 83.2%; Pred. No. 2.1e-154;
Matches 543; Conservative 40; Mismatches 39; Indels 31; Gaps 7;
3 LWTVYVGVVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPNQVLENTENF 62
28 LWTVYVGVVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPNQVLENTENF 87
63 NMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELISIVVWEO 122
88 NMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELISIVVWEO 141
123 RKGEMRNCSPNITSIRKQVOREYALFYKLDVEPDDDKNNTNTTKYRLINCNTSVITQ 182
142 MDRGEIKNCSFKVTTISIRNMQKEXALFYKLDVPIDN-----DNTSYKLINCNTSVITQ 196
183 ACPKVSFEPIHYCTPTGTFALLKNDKFKNGTCTNVTSTVQCTHGIPTVSTOLLING 242
197 ACPKVSFEPIHYCAPAGFALKNDKFKNGSGTCTNVTSTVQCTHGIPTVSTOLLING 256
243 SLAEWVIRSENFNTNAKTIIVQLNVSVVEINCTRENHTRKRVTLGPRVWVTGIELG 302
257 SLAEWVIRSENFNTNAKTIIVQLKESVEINCTRENHTRKRVTLGPRVWVTGIELG 316
303 NTRQAHNISRAQWNTLQOIATTLREQGNTKIAPNQSSGSDPEIVMHSFNGGGEFFYC 362
317 DIRQAHNISGEGKWNNTLQVITKLQAQFGNKTIIVFKQSSGSDPEIVMHSFNGGGEFFYC 376
363 NSTQLFNSAWNT-----SNGTWSVTRKQKGDITLPCRKQIINRQVGVKAWYALP 417
377 NSTQLFNSAWNTIGPNTNGT-----ITLPCRKQIINRQVGVKAWYALP 423
418 IKGLRCSNITGLLTRDGGGE-NOTTEIFRPGGDMRDNRSLEYKYKVKIPLGVA 476
424 IRSQIRCSNITGLLTRDGGGEISNTTEIFRPGGDMRDNRSLEYKYKVKIPLGVA 483
477 PTKAKRRVVOREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVQARQLLSGIVQOQNLL 536
484 PTKAKRRVVOREKRAV-TLGAMFLGFLGAAGSTMGARSITLTVQARQLLSGIVQOQNLL 542

QY 537 RAIKAQOHLQLTWGKIQOARILAVERYLKQQLLGFWGSGKLICTTAVPWNASWN 596
Db 543 RAIKAQOHLQLTWGKIQOARILAVERYLKQQLLGFWGSGKLICTTAVPWNASWN 602
QY 597 KTLDDQIWNNTWMEWDREIDNTHLYTTLIESONQOEKQQLQLDKWASL 649
Db 603 KSLDDQIWNNTWMEWDREIDNTHLYTTLIESONQOEKQQLQLDKWASL 655

RESULT 6

ABB06211
ID ABB06211 standard; Protein; 842 AA.
XX
XX ABB06211;
AC
XX
XX 15-MAY-2002 (first entry)
DT
XX
XX HIV Env isolate SF162 amino acid sequence.
DE
XX
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy.
XX
XX Human immunodeficiency virus type C.
OS Synthetic.
XX
XX WO200204493-A2.
PN
XX
XX 17-JAN-2002.
PD
XX
XX 05-JUL-2001; 2001WO-US21241.
PF
XX
XX 05-JUL-2000; 2000US-0610313.
PR
XX
XX (CHIR) CHIRON CORP.
PA
XX
XX (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
PI
XX
XX WPI; 2002-154920/20.
DR
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -
XX
XX Disclosure; Fig 105; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (I). (i) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 842 AA;

Query Match 82.6%; Score 2870.5; DB 23; Length 842;
Best Local Similarity 83.0%; Pred. No. 5.8e-154;
Matches 542; Conservative 39; Mismatches 41; Indels 31; Gaps 7;

QY 3 LWTVYVGVVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPNQVLENTENF 62
Db 28 LWTVYVGVVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPNQVLENTENF 87
QY 63 NMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELISIVVWEO 122
Db 88 NMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTKGSN----WKE 141

DE Human Immunodeficiency Virus-1 strain BA-L envelope protein.
 XX Acquired immune deficiency syndrome; AIDS; envelope protein;
 KW env gene; vaccine.
 XX
 OS Human immunodeficiency virus type 1 (strain BA-L).
 XX
 PN US5576000-A.
 XX
 PD 19-NOV-1996.
 XX
 XX 15-FEB-1995; 95US-0388809.
 XX
 PF 17-OCT-1990; 90US-0599491.
 PR 25-FEB-1993; 93US-0022835.
 PR 15-FEB-1995; 95US-0388809.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
 PI Popovic M, Reitz MS;
 XX
 DR WPI; 1997-011206/01.
 DR N-PSDB; AAT58551.
 XX
 XX New isolated envelope protein of HIV-1 strain BA-L and recombinant
 PT equivalents - useful as immunogens for vaccines and antibody prodn.,
 PT typical of US clinical isolates
 XX
 XX Claim 1; Fig 9; 86pp; English.
 XX
 CC A HindIII fragment of unintegrated viral DNA representing the BA-L
 CC genome was cloned into lambda phage Charon 28 DNA from total DNA of
 CC peripheral blood lymphocytes infected with and producing HIV-1 (BA-L).
 CC A positive clone was selected by hybridisation using a HIV-1 env
 CC probe. This clone, designated BA-L1, was found to contain the
 CC entire gene for the envelope protein on a 2.8 kb HindIII-XbaI
 CC fragment and a 0.4 kb EcoRI-HindIII fragment. When cloned together
 CC these fragments comprise the env gene, as well as the coding regions
 CC for rev and the rev-responsive element of env, both necessary for
 CC efficient expression in eukaryotic cells. The claimed recombinantly
 CC produced envelope protein can be used as an immunogen for raising
 CC antibodies against HIV.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Sequence 855 AA;
 SQ

Query Match 81.6%; Score 2834; DB 18; Length 855;
 Best Local Similarity 82.2%; Pred.No. 6.7e-152;
 Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;
 3 LMVTYGVGVVKEATTTLCASDRKAYDTEVENVWATHACVPTDENPQVELKNVTENF 62
 33 LMVTYGVGVVKEATTTLCASDRKAYDTEVENVWATHACVPTDENPQVELKNVTENF 92
 63 NMKNMNVQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLN--TN-NTNTTELSIIV 119
 93 NMKNMNVQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNATNGNDNTTSSRGMV 152
 120 WEQRGKGMNRGCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYLINCNTSV 179
 153 ----GGGEMKNCNFNTTWIRGKQVEYALFYKLDIAPIDNN----SNNRYLISNTSV 204
 180 ITQACPKVFEPIPIHYCTPTGTFALLKCNCKFNGTGPCNTNVSTVQCTHGIRPVSTQLL 239
 205 ITQACPKVFEPIPIHYCAPAGAILKCKCKFNGKGPCTNVSTVQCTHGIRPVSTQLL 264
 240 LNSGLAEEVIRSENFNTNAKTIIVQLNVSVIEINCTRNPNHTRKRVTLGPRVWTTGE 299
 265 LNSGLAEEVIRSENFNTNAKTIIVQLNVSVIEINCTRNPNHTRKRVTLGPRVWTTGE 324
 300 ILGNIRCAHNTSRAOWNTLQOIAITLREOFNKTIAFNQSSGGDPRIWHSFNCGGEF 359

Db 325 IIGDIRQAHCLNSRAKWNNTLNKIVIKLREQFGNKTIIVFKHSGGDEIVTHSFNCGGEF 384
 QY 360 FYCNSQLFNSAWNTVSTNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAMYALPIK 419
 Db 385 FYCNSQLFNSAWNTVSTNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAMYALPIK 438
 QY 420 GLIRCSSNITGLLLTRDGGGNGQTTEIPRPGGDMRDNRSLEYKYKVIKIEPLGVAPTK 479
 Db 439 GQIRCSSNITGLLLTRDGGGNGQTTEIPRPGGDMRDNRSLEYKYKVIKIEPLGVAPTK 498
 QY 480 AKRRVVQREKRAVGMLGAMFLGLGAAGSTMATVQARQLLSGIVQQNNLLRAI 539
 Db 499 AKRRVVQREKRAVG-IGAVFLGLGAAGSTMATVQARQLLSGIVQQNNLLRAI 557
 QY 540 KAAQHLLQLTVMGIKQLQARILAVERYLKDQQLLGPWCGSKLICITTAVPWNASWNSKTL 599
 Db 558 EAAQHLLQLTVMGIKQLQARILAVERYLKDQQLLGPWCGSKLICITTAVPWNASWNSKSL 617
 QY 600 DOIWNMTMWEDREIDNYTHLYTLIBESQNOQKNOQELIQLDKWASL 649
 Db 618 NKIWDNMTWIRWDREINNYTTHYSLIESQNOQKNOQELIQLDKWASL 667

RESULT 9
 AAW88113
 ID AAW88113 standard; Protein; 855 AA.
 XX AC AAW88113;
 XX DT 09-APR-1999 (first entry)
 XX DE Env protein of the BA-L strain of Human immunodeficiency virus type 1.
 XX KW HIV-1; HIV-1 strain BA-L; env protein; vaccine;
 XX OS Human immunodeficiency virus type 1.
 XX Key Location/Qualifiers
 FT Misc-difference 11 /note= "Gln encoded by CG"
 FT
 FT
 PN US5869313-A.
 XX 09-FEB-1999.
 XX 14-MAY-1996; 96US-0647714.
 XX 17-OCT-1990; 90US-0599491.
 XX 25-FEB-1993; 93US-0022835.
 XX 15-FEB-1995; 95US-0388809.
 XX 14-MAY-1996; 96US-0647714.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Franchini G, Gallo RC, Gartner S, Lori FC, Markham PD;
 PI Popovic M, Reitz MS;
 XX WPI; 1999-152779/13.
 DR N-PSDB; AAX04767.
 XX
 XX DNA encoding env protein of the human immune deficiency virus
 PT isolate BA-L - useful for producing protein for use in vaccines, as
 PT assay reagent and to generate antibodies
 XX
 XX Example 1; Fig 9A-C; 87pp; English.
 CC The present sequence represents the envelope protein of the BA-L
 CC (ATCC 40890) strain of Human immunodeficiency virus type 1 (HIV-1)
 CC strain MN-ST1. BA-L is more typical of United States isolates of
 CC HIV-1 than previously known strains. Recombinant, complete env protein
 CC of the BA-L strain is used as a vaccine component and for immunotherapy
 CC of existing HIV infections, to detect HIV-specific antibodies, e.g. in

```
CC donated blood, and as an immunogen to raise specific antibodies, for
CC HIV-1 diagnosis.
XX
SQ Sequence 855 AA;

Query Match      81.6%; Score 2834; DB 20; Length 855;
Best Local Similarity 82.2%; Pred. No. 6.7e-152;
Matches 534; Conservative 45; Mismatches 53; Indels 19; Gaps 6;

QY 3 LWVTYYGVVWKEATTTLCASDAKAYDTEAHNWWATHACVPTNPQEVVLENTENF 62
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 LWVTYYGVVWKEATTTLCASDAKAYDTEAHNWWATHACVPTNPQEVVLENTENF 92
QY 63 NNMKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLN--TN--NTNTTSLIIV 119
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 NNMKNMVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNATNGDNTTSSRGV 152
QY 120 WEQRKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTKYELINCVTSV 179
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 ----GGGEMKNCSEFNITTSIRDKVQREYALFYKLDIAPIDNN---SNNRYELISCVTSV 204
QY 180 ITQACPKVSFEPIPIHYCTPTGFPALLKNDKFKNGTGCTNVSTVQCTHGIRPVVSTOLL 239
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
205 ITQACPKVSFEPIPIHYCAPAGFAILCKDKKFKNGTGCTNVSTVQCTHGIRPVVSTOLL 264
QY 240 LNSGLAEBEVVIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRKRVTLGPGRVWYTTGE 299
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 LNSGLAEBEVVIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRKRVTLGPGRVWYTTGE 324
QY 300 ILGNTRQAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEFFYC 359
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 IIGDIRQAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEFFYC 384
QY 360 FYCNSTQLFNSAWNTSNGTWVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPIK 419
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 FYCNSTQLFNSAWNTSNGTWVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPIK 438
QY 420 GLIRCSSNITGLLTFRDGGGENTTEIFRPGGDMRDNRSLYKYVVKVKEIPLGVAPT 479
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
439 GQIRCSSNITGLLTFRDGGGENTTEIFRPGGDMRDNRSLYKYVVKVKEIPLGVAPT 498
QY 480 AKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMATVQARQLLSGIVQOQNLLRAI 539
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
499 AKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMATVQARQLLSGIVQOQNLLRAI 557
QY 540 KAQHLQLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPNWASNSKTL 599
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
558 EAQHLQLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPNWASNSKTL 617
QY 600 DQIWNNTWMEWDREIDNTYLIYTLIBESQNOQKQELQLDKWASL 649
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
618 NKIWNMTWIEWDREINNTYLIYTLIBESQNOQKQELQLDKWASL 667

RESULT 10
AAB61505
ID AAB61505 standard; Protein; 643 AA.
XX
AC AAB61505;
XX
DT 05-APR-2001 (first entry)
XX
DE HIV-1 SOS gp140 glycoprotein.
XX
KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;
KW gp140.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200100648-A1.
XX
PD 04-JAN-2001.
XX
```

525 QRMQLQTVMGIKQARVLAVERYLGQQLGIWGCCKLICCTTAVPWNASWSNKSILDR1 584

603 WNNMTWMEWDREIDNTHLYTLIESQNOEKQOELLQDKWASL 649

585 WNNMTWMEWDREIDNTHLYTLIESQNOEKQOELLQDKWASL 631

RESULT 11

AAP80967

ID AAP80967 standard; protein; 851 AA.

XX AC

XX AAP80967;

DT 25-MAR-2003 (updated)

DT 15-NOV-1990 (first entry)

XX HIV protein HT6.

XX HIV; HT6; gp160; envelope protein; RF; AIDS.

XX Human immunodeficiency virus variant RF.

XX BP272858-A.

XX 29-JUN-1988.

XX 14-DEC-1987; 87EP-0310967.

XX 31-AUG-1987; 87US-0091481.

PR 15-DEC-1986; 86US-0941111.

XX (REPK) REPLIGEN CORP.

XX Rusche J, Lynn D, Carson H, Putney S, Jellis CL;

XX WPI; 1988-176944/26.

DR N-PSDB; AAN80948.

XX Prodn. of recombinant HIV envelope proteins in insect cells -

PT useful as vaccine against AIDS and for diagnosis and therapy.

XX Disclosure; Page ?; 4pp; English.

XX The sequence is the result of cloning a hybrid envelope gene from

CC HIV variants BH10 and RF. A central portion of the RF gene was

CC used, the rest being from te distantly related variant BH10. The

CC resulting clone, pACHR6 produces a hybrid gp 160 envelope protein

CC with novel immunological and antigenic characteristics. It may

CC be used to as a vaccine and for diagnosis and therapy of AIDS.

CC See also AAP80966.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 851 AA;

SQ Query Match 81.1%; Score 2816; DB 9; Length 851;

Best Local Similarity 82.0%; Pred. No. 6.9e-151;

Matches 533; Conservative 42; Mismatches 53; Indels 22; Gaps 7;

3 LWVTVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVTFN 62

34 LWVTVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVTFN 93

63 NNMKNMVEQMEHDIISLDQSLKPCVKLTPLCVLNCITDLNNTNTTELSIIIVWEQ 122

94 NNMKNMVEQMEHDIISLDQSLKPCVKLTPLCVLNCITDLNNTNTSSSGRIME- 150

123 RGKGMNRCNFNITTSIRDKVOREYALFVKLDVERIDDKNTNNTKYELINCNTSVITQ 182

151 --XGEIKNCFNMTSIRKQVEYAFVKLDIIPIDN-----DTTSYTLTSCNTSVITQ 203

183 ACPKVSFEPIHYCTPTGALLKCNKDKFNFGTGPCTNVSTVQCTHGIRPVYSTOLLNG 242

204 ACPKVSFEPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVQCTHGIRPVYSTOLLNG 263

QY 243 SLABEEVIRSENFTNAKTIIVQLNVSVIEINCTRPNNHTRKRVTLGPGRVVTTSEILG 302

DB 264 SLABEEVIRSENFTNAKTIIVQLNVSVIEINCTRPNNHTRKRVTLGPGRVVTTSEILG 323

QY 303 NIRQAHCNISRAQWNTTLOQIATTLREOFNGKTIAFNCSGGDPEIVMHSFNCGBEFYC 362

DB 324 DIRKAHCNLSRAQWNTTLOQIATTLREOFNGKTIAFNCSGGDPEIVMHSFNCGBEFYC 383

QY 363 NSTQLFNSAWNVT--SNGTWSVTRKOKDGTCDITLPCRIKQIINRWQVVGKAMVYALPIKG 420

DB 384 NTTQLFNSAWNVT--SNGTWSVTRKOKDGTCDITLPCRIKQIINRWQVVGKAMVYALPIKG 435

QY 421 LIRCSSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTK 479

DB 436 QIKCISNITGLLLTRDGGGE-TNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTK 495

QY 480 AKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRAI 539

DB 496 AKRRVVQREKRAVG-IGALFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRAI 554

QY 540 KAOQHLLQTVMGIKQARVLAVERYLKDQQLLGFWGCCKLICCTTAVPWNASWSNKSIL 599

DB 555 EAOQHLLQTVMGIKQARVLAVERYLKDQQLLGFWGCCKLICCTTAVPWNASWSNKSIL 614

QY 600 DQIWNMTWMEWDREIDNTHLYTLIESQNOEKQOELLQDKWASL 649

DB 615 EQIWNMTWMEWDREIDNTHLYTLIESQNOEKQOELLQDKWASL 664

RESULT 12

AAM53112

ID AAM53112 standard; Protein; 855 AA.

XX AC

XX AAM53112;

DT 25-JUN-1998 (first entry)

XX ENV protein contained in a complete ARV-2 nucleotide sequence.

DE ARV-2; enhanced promoter; gene expression; cytomegalovirus;

KW HIV; AIDS; ENV protein.

XX Human immunodeficiency virus type 1.

OS US5688688-A.

PN 18-NOV-1997.

XX 10-AUG-1994; 94US-0288336.

XX 24-DEC-1987; 87US-0138894.

PR 31-OCT-1984; 84US-0667501.

PR 30-JAN-1985; 85US-0696534.

PR 06-SEP-1985; 85US-0773447.

PR 17-AUG-1992; 92US-0931191.

PR 28-JUN-1993; 93US-0083391.

PR 17-AUG-1993; 93US-0107377.

PR 10-AUG-1994; 94US-0288336.

XX (CHIR) CHIRON CORP.

XX Chapman BS, Dina D, Haigwood NL, Luciw PA, Rosenberg S;

PI Thayer RM;

XX WPI; 1998-007982/01.

DR N-PSDB; AAV04733.

XX Enhanced promoter for gene expression - comprising cytomegalovirus

PT immediate early promoter plus intron

XX Example 1; Fig 4C-P; 9pp; English.

CC This sequence represents the ENV protein contained in a complete
 CC nucleotide sequence of ARV-2 derived from partial sequences of several
 CC ARV clones. The invention provides a method for construction of a vector
 CC for expression of a polypeptide in a mammalian cell, comprising a
 CC polypeptide coding sequence operably linked downstream of an enhanced
 CC promoter. The enhanced promoter comprises the human cytomegalovirus
 CC immediate early region (HCMV IE1) promoter and the first intron proximate
 CC to the 3' end of the HCMV IE1 promoter. The polypeptide can be any of the
 CC HIV recombinant polypeptides and especially HIV gp120. Expression of HIV
 CC gp120 by COS 7 cells transfected with pCMV6a containing the gp120 coding
 CC region, where pCMV6a is a vector containing the above enhanced promoter,
 CC is increased by a factor of 50-100 compared with the use of a vector
 CC containing the SV40 early promoter.

XX Sequence 855 AA;

Query Match 81.0%; Score 2815.5; DB 19; Length 855;
 Best Local Similarity 81.2%; Pred. No. 7.4e-151;
 Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;

QY 3 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 62
 DB 33 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 92
 QY 63 NWWKNNVVEQMHEDTISLWDSQKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVWVEQ 122
 DB 93 NWWKNNVVEQMHEDTISLWDSQKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146
 QY 123 RGKGMRCNSFNITTSIRDKVQREYALFKLDVEPIDDNKNTNTNTKYLINCNTSVITQ 182
 DB 147 EIKGEIKNCSEFNITTSIRDKIKENALFRNLDVVDIDNASTTNTVNYRLIHCRNSVITQ 206
 QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKKFNGTGCTNVSTVQCHGIRPVVSTOLLNG 242
 DB 207 ACPKVSFPIPIHYCTPTGPFALLKNDKKFNGTGCTNVSTVQCHGIRPVVSTOLLNG 266
 QY 243 SLAEVEVIRSENFNTNAKTIIVQLNVSEINCTRNPHHTRKRVTLGPRVWYTTGEILG 302
 DB 267 SLAEVEVIRSENFNTNAKTIIVQLNVSEINCTRNPHHTRKRVTLGPRVWYTTGEILG 326
 QY 303 NTRQAHCNISRAQWNTLQIATTLREQFG-NKTIAPNQSSGGDPEIVMHSFNCGEFFY 361
 DB 327 DIRKAHCNISRAQWNTLQIATTLREQFG-NKTIAPNQSSGGDPEIVMHSFNCGEFFY 386
 QY 362 CNSTOLFNSANVVT-SNGTWSVTRKQKDTGDIITLPCRIKQILNQRVQVGMALPIKG 420
 DB 387 CNTTOLFNTWRLNHTGT-----KNDT---IILPCRIKQILNQRVQVGMALPIKG 438
 QY 421 LIRCSSNITGLLLTRDGG-GENQTTEIFRPGGDMRDNRSELYKVKVVKTEPLGVAPTK 479
 DB 439 QIQCSSNITGLLLTRDGGNTVNTDTEVFRPGGDMRDNRSELYKVKVVKTEPLGVAPTK 498
 QY 480 AKRRVVQREKRAVGMGLGFLGAAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAI 539
 DB 499 AKRRVVQREKRAVGMGLGFLGAAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAI 558
 QY 540 KAOQHLLQITWGIKQIARILAVRYLKDOQLLGFWGCSSGKLICTTAVPNWASNKTL 599
 DB 559 EAOQHLLQITWGIKQIARILAVRYLKDOQLLGFWGCSSGKLICTTAVPNWASNKSL 618
 QY 600 DOIWNNTWMEWDREIDNVTIYTLIYESQNOQKQOELLQDKWASL 649
 DB 619 EDIWNNTWMEWDREIDNVTIYTLIYESQNOQKQOELLQDKWASL 668

RESULT 13

AA77298

ID AA77298 standard; Protein; 855 AA.

XX AA77298;

AC AA77298;

XX 22-MAY-2000 (first entry)

DE HIV-1 (ATCC CRL 8597) env protein.
 XX HIV-1: immunoassay; antigen; pol fragment; p31; recombinant; antibody;
 KW detection; env protein.
 XX Human immunodeficiency virus type 1 'ATCC CRL 8597'.
 XX US6013432-A.
 XX 11-JAN-2000.
 XX 17-MAY-1995; 95US-0443434.
 PR 08-JUL-1993; 93US-00899407.
 PR 24-DEC-1987; 87US-0138894.
 PR 17-AUG-1992; 92US-0951154.
 PR 31-OCT-1984; 84US-0667501.
 PR 30-JAN-1985; 85US-0696534.
 PR 06-SEP-1985; 85US-0773447.
 XX (CHIR) CHIRON CORP.
 PA Luciw PA, Dina D;
 PI WPI; 2000-170256/15.
 DR N-PSDB; AAZ90201.
 XX Immunoassay for antibodies against human immune deficiency virus, for
 PT diagnosing infection, uses an immunogenic fragment of the pol protein
 PT as antigen -
 XX Example 1; Fig 4K-O; 99pp; English.
 XX The invention relates to the improvement of HIV-1 immunoassays by the
 CC use of an HIV-1 antigen comprising an immunogenic fragment of
 CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately
 CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an
 CC NdeI site at position 5131 of the genome (the proviral DNA sequence is
 CC given in AAZ90201). The immunogenic pol fragment is not immunologically
 CC cross-reactive with human T cell lymphotropic viruses I or II. The
 CC invention also encompasses the use of p31 as an antigen. The recombinant
 CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or
 CC in mammalian cells. Immunoassays using the recombinant HIV proteins may
 CC be used to diagnose and stage HIV-1 infections. Sequences
 CC AA77294-77299 represent proteins encoded by the genome of HIV-1 (ATCC
 CC CRL 8597).

XX Sequence 855 AA;

Query Match 81.0%; Score 2815.5; DB 21; Length 855;
 Best Local Similarity 81.2%; Pred. No. 7.4e-151;
 Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;

QY 3 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 62
 DB 33 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 92
 QY 63 NWWKNNVVEQMHEDTISLWDSQKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVWVEQ 122
 DB 93 NWWKNNVVEQMHEDTISLWDSQKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146
 QY 123 RGKGMRCNSFNITTSIRDKVQREYALFKLDVEPIDDNKNTNTNTKYLINCNTSVITQ 182
 DB 147 EIKGEIKNCSEFNITTSIRDKIKENALFRNLDVVDIDNASTTNTVNYRLIHCRNSVITQ 206
 QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKKFNGTGCTNVSTVQCHGIRPVVSTOLLNG 242
 DB 207 ACPKVSFPIPIHYCTPTGPFALLKNDKKFNGTGCTNVSTVQCHGIRPVVSTOLLNG 266
 QY 243 SLAEVEVIRSENFNTNAKTIIVQLNVSEINCTRNPHHTRKRVTLGPRVWYTTGEILG 302
 DB 267 SLAEVEVIRSENFNTNAKTIIVQLNVSEINCTRNPHHTRKRVTLGPRVWYTTGEILG 326

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:09:27 ; Search time 22 Seconds
(without alignments)
1248.170 Million cell updates/sec

Title: US-09-938-406-1_COPY_33_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVWKEATT.....QNQKEKNOQLLDKWSL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	82.1	850	2	US-08-448-603A-28
2	2852	82.1	850	3	US-09-134-075-28
3	2852	82.1	850	4	US-09-492-739-28
4	2834	81.6	855	1	US-08-022-835-6
5	2834	81.6	855	1	US-08-388-803-6
6	2834	81.6	855	2	US-08-647-714-6
7	2818	81.1	855	3	US-07-956-483-14
8	2814.5	81.0	826	1	US-08-375-510-2
9	2814.5	81.0	826	2	US-08-487-657-2
10	2814.5	81.0	854	4	US-09-309-572-23
11	2814	81.0	880	2	US-08-788-835-7
12	2814	81.0	880	3	US-09-157-963-7
13	2809.5	80.9	856	2	US-07-916-098A-2
14	2809.5	80.9	857	2	US-08-448-603A-30
15	2809.5	80.9	857	3	US-09-134-075-30
16	2809.5	80.9	857	4	US-09-492-739-30
17	2807.5	80.8	855	3	US-07-956-483-15
18	2806.5	80.8	856	3	US-09-124-900-9
19	2806.5	80.8	863	3	US-08-463-210-11
20	2795.5	80.5	856	4	US-09-337-387-11
21	2795	80.5	861	1	US-08-127-499A-14
22	2795	80.5	861	1	US-08-482-847-14
23	2795	80.5	861	3	US-07-956-483-10
24	2795	80.5	861	3	US-08-472-240A-1
25	2795	80.5	861	3	US-08-472-240A-7
26	2795	80.5	861	4	US-08-817-441-103
27	2787	80.2	861	3	US-07-956-483-16

Sequence 13, Appl
Sequence 5, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 2, Appl
Sequence 10, Appl
Sequence 6, Appl
Sequence 3, Appl
Sequence 12, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 11, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-448-603A-28
; Sequence 28, Application US/08448603A
; Patent No. 5864027
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448.603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/072,833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Halliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-448-603A-28

Query Match 82.1%; Score 2852; DB 2; Length 850;

Best Local Similarity 83.0%; Pred. No. 2.4e-227;

Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;

3 LWTVYGVVWKEATTTLFCASAKAYDTEAHNVWATHACVPTNPQEVVLENTFNF 62

Db 33 LWTVVYGVFWKEATTTLCASDAKAYDETEHNVWATHACVPTDNPQOELGVNTENF 92
Qy 63 NMKNVVEQVHEDIIISLWQSLKPCVKLPLCVTLNCTDLNNTNTNTTNTSIIIVWEQ 122
Db 93 NMKNVVEQVHEDIIISLWQSLKPCVKLPLCVTLNCTDL--KNAINTTSS--WGK 146
Qy 123 RGKEMENCSFNITTSIRDKVOREYALFYKLDVPEIDDKNTNTNTKYRLINCMNTSITO 182
Db 147 MERGEIKNCFSNVTSIRDKMKNEYALFYKLDVVPIDN-----DNTSYRLISCMNTSITO 201
Qy 183 ACQVSEPIPIHYCTPTGFALLKCNKFGTGPCINVTQCTHGRIPVSTQLLING 242
Db 202 ACQVSEPIPIHYCAPAGFALLKCRDKKFGTGPCINVTQCTHGRIPVSTQLLING 261
Qy 243 SLABEEVIRSENTNNAKTIIVQVNSVEINCTRPNNHTRKRVLTGPRVWYTTGEILG 302
Db 262 SLABEEVIRSENFSDNAKTIIVQVNSVEINCTRPNNHTRRSIHIGGRAFYATGEIIG 321
Qy 303 NIRAHCNI SPAQNNLTQIATTLRQFGNKTIATFQSGGDDPEIIVMHSFNCGEFFYC 362
Db 322 DIRQAHCLNSTKWNLTQIATTLRQFGNKTIATFQSGGDDPEIIVMHSFNCGEFFYC 380
Qy 363 NSTOLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYPKGLI 422
Db 381 NTFPLFNSWNTY--TWNTGSDTGRNITLQCRKQIINRWQVVGKAMYPKGLI 438
Qy 423 RCSSNITGLLTRDGGGENTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 482
Db 439 RCSSNITGLLTRD--GGNNSSETEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 497
Qy 483 RVQREKRAVGLGAMFLGFLGAAGSTMGATSMALTVOAROLLGSIQOQNNLLRAIKAO 542
Db 498 RVQREKRAVG--IGAVFLGFLGAAGSTMGASVLTVOAROLLGSIQOQNNLLRAIAE 556
Qy 543 QHLLQLTWGIKQIARILAVERYLKQOQLLGFWSGCKLICCTTAVPWNASWKNLTDOI 602
Db 557 QHLLQLTWGIKQIARILAVERYLKQOQLLGIWCSGCKLICCTTAVPWNASWKNLSLDKI 616

RESULT 2

US-09-134-075-28
; Sequence 28, Application US/09134075
; Patent No. 6042836
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,075
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,603
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haliday, Emily

; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 850 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-134-075-28

Query Match 82.1%; Score 2852; DB 3; Length 850;
Best Local Similarity 83.0%; Pred. No. 2.4e-227;
Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;

Qy 3 LWTVVYGVFWKEATTTLCASDAKAYDETEHNVWATHACVPTDNPQOELGVNTENF 62
Db 33 LWTVVYGVFWKEATTTLCASDAKAYDETEHNVWATHACVPTDNPQOELGVNTENF 92
Qy 63 NMKNVVEQVHEDIIISLWQSLKPCVKLPLCVTLNCTDLNNTNTNTTNTSIIIVWEQ 122
Db 93 NMKNVVEQVHEDIIISLWQSLKPCVKLPLCVTLNCTDL--KNAINTTSS--WGK 146
Qy 123 RGKEMENCSFNITTSIRDKVOREYALFYKLDVPEIDDKNTNTNTKYRLINCMNTSITO 182
Db 147 MERGEIKNCFSNVTSIRDKMKNEYALFYKLDVVPIDN-----DNTSYRLISCMNTSITO 201
Qy 183 ACQVSEPIPIHYCTPTGFALLKCNKFGTGPCINVTQCTHGRIPVSTQLLING 242
Db 202 ACQVSEPIPIHYCAPAGFALLKCRDKKFGTGPCINVTQCTHGRIPVSTQLLING 261
Qy 243 SLABEEVIRSENTNNAKTIIVQVNSVEINCTRPNNHTRKRVLTGPRVWYTTGEILG 302
Db 262 SLABEEVIRSENFSDNAKTIIVQVNSVEINCTRPNNHTRRSIHIGGRAFYATGEIIG 321
Qy 303 NIRAHCNI SPAQNNLTQIATTLRQFGNKTIATFQSGGDDPEIIVMHSFNCGEFFYC 362
Db 322 DIRQAHCLNSTKWNLTQIATTLRQFGNKTIATFQSGGDDPEIIVMHSFNCGEFFYC 380
Qy 363 NSTOLFNSAMVNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYPKGLI 422
Db 381 NTFPLFNSWNTY--TWNTGSDTGRNITLQCRKQIINRWQVVGKAMYPKGLI 438
Qy 423 RCSSNITGLLTRDGGGENTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 482
Db 439 RCSSNITGLLTRD--GGNNSSETEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 497
Qy 483 RVQREKRAVGLGAMFLGFLGAAGSTMGATSMALTVOAROLLGSIQOQNNLLRAIKAO 542
Db 498 RVQREKRAVG--IGAVFLGFLGAAGSTMGASVLTVOAROLLGSIQOQNNLLRAIAE 556
Qy 543 QHLLQLTWGIKQIARILAVERYLKQOQLLGFWSGCKLICCTTAVPWNASWKNLTDOI 602
Db 557 QHLLQLTWGIKQIARILAVERYLKQOQLLGIWCSGCKLICCTTAVPWNASWKNLSLDKI 616
Qy 603 WNNMTWMEWDREIDNYTHLIYTLIEESQOQEKQOELLQDKWASL 649
Db 617 WNNMTWMEWDREIDNYTHLIYTLIEESQOQEKQOELLQDKWASL 663

RESULT 3

US-09-492-739-28
; Sequence 28, Application US/09492739
; Patent No. 6331404
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:

ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
STREET: 3 Embarcadero Center
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/492,739
FILING DATE: 27-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Haliday, Emily
REGISTRATION NUMBER: 38903
REFERENCE/DOCKET NUMBER: 14918-704
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-393-2000
TELEFAX: 415-393-2286
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 850 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-492-739-28

Query Match 82.1%; Score 2852; DB 4; Length 850;
Best Local Similarity 83.0%; Pred. No. 2.4e-227;
Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;
3 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62
33 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92
63 NWWKNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNTELSIIWVEQ 122
93 NWWKNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLN--KNATNTSSS---WKG 146
123 RKGEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNNTNNTKYRLNCTNSVITQ 182
147 MERGEIKNCSPNITTSIRDKMNEALFYKLDVPIDN-----DNSTYELISCTNSVITQ 201
183 ACPKVSFEPIPIHYCTPTGTFALLKNDKFGNGTGPCTNVSTVQCTHGRIPVYSTOLLNG 242
202 ACPKVSFEPIPIHYCAPAGFALLKCRDKKFGNGTGPCTNVSTVQCTHGRIPVYSTOLLNG 261
243 SLAEVEVIRSENFNTNAKTIIVQLNVSVEINCTPNNHTRKRVTLGPGRVWYTTGEILG 302
262 SLAEVEVIRSENFNTNAKTIIVQLNVSVEINCTPNNHTRKRVTLGPGRVWYTTGEILG 321
303 NTRQAHNCISRAQWNNLTLOQIATTLREQGNKTIAPNQSOGDPEIVMHSFNCGGEFFYC 362
322 DIRQAHNCISRAQWNNLTLOQIATTLREH-NTIVFNHSSGGDPEIVMHSFNCGGEFFYC 380
363 NQTQIFNSAWNTSNGTSVTRKQKDTGDIITLPCRKQIINRWQVGVGKAMYPALPIKGLI 422
381 NTPPLFNSWNTVY--TWNTSGSDTGRNITLQCKRIKQIINWQEVGKAMYPALPIRGI 438
423 RCSSNITGLLLTRDGGNGNQTTTEIRPGGGRDNRWSELYKYKVKIPLGVATTKAKR 482
439 RCSSNITGLLLTRD--GGNNSSETTEIRPGGGRDNRWSELYKYKVKIPLGVATTKAKR 497
483 RVVQREKRAVGMGLGFLGAGSTMGATSMALTVOARQLLSGIVQOONLLRAIKAO 542

498 RVWQREKRAVG-IGAVFLGFLGAGSTWGAASVTITVQARLLSGIVQOONLLRAIEAE 556
543 QHLLQLTWGIKQLQARILAVERYLKDQQLLGFWGCGSKLICTTAVPMNASWSNKTLDQI 602
557 QHLLQLTWGIKQLQARVLAVERYLKDQQLLGFWGCGSKLICTTAVPMNASWSNKSLDKI 616
603 WNNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLQLDKWASL 649
617 WDNMTWMEWEIEDNYTSLIYSLIESQNOQEKNEQELLELDKWASL 663
RESULT 4
US-08-022-835-6
Sequence 6, Application US/08022835
Patent No. 5420030
GENERAL INFORMATION:
APPLICANT: Reitz Jr., Marvin S.
APPLICANT: Franchini, Genoveffa
APPLICANT: Markham, Phillip D.
APPLICANT: Gallo, Robert C.
APPLICANT: Lori, Franco C.
APPLICANT: Popovic, Mikulas
APPLICANT: Gartner, Suzanne
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: CUSHMAN, DARBY & CUSHMAN
STREET: Eleventh Floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-022-835-6

Query Match 81.6%; Score 2834; DB 1; Length 855;
Best Local Similarity 82.2%; Pred. No. 7.5e-226;
Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;
3 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62
33 LWVTYVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92
63 NWWKNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLN--TN--NTNTTSLIIV 119
93 NWWKNNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLNATNGDNTTSSRGW 152

CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM: DISK, 3.5"
 MEDIUM TYPE: FLOPPY DISK, 3.5"
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/647,714
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/388,809
 FILING DATE: 15-FEB-1995
 APPLICATION NUMBER: US/08/022,835
 FILING DATE: 25-FEB-1993
 APPLICATION NUMBER: US/07/599,491
 FILING DATE: 17-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: LESLIE A. SERUNIAN
 REGISTRATION NUMBER: 35,353
 REFERENCE/DOCKET NUMBER: 2026-4092US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 855 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-647-714-6

Query Match 81.6%; Score 2834; DB 2; Length 855;
 Best Local Similarity 82.2%; Pred. No. 7.5e-226;
 Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;
 QY 3 LWVTYVGVVWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
 Db 33 LWVTYVGVVWKEATTTLCFASDRKAYDTEVHNWATHACVPTNPQEVVLENTNF 92
 QY 63 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLN-TN-NTNTTTELSIIVV 119
 Db 93 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNATNGNDTNTSSRGWV 152
 QY 120 WEQKQKGMNRCFNITTSIRKQVREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSV 179
 Db 153 ----GGGEMKNGSFNITIRGKVQKEYALFYKLDIADIPDN---SNNRYRLISCNTSV 204
 QY 180 ITQACPVSFEPIPHYCTPTFALLKNDKKFNGTGPCTNVSTVQCTGIRPVVSTQLL 239
 Db 205 ITQACPVSFEPIPHYCAPAGFALLKNDKKFNGTGPCTNVSTVQCTGIRPVVSTQLL 264
 QY 240 LNSGLAEVEVIRSNFTNNAKTIIVQLNVSVVEINCTRNHNTRKRVLTGPGRVVYTCGE 299
 Db 265 LNSGLAEVEVIRSNFADNAKVIIVQLNVSVVEINCTRNHNTRKRSIHIGPGRVYTCGE 324
 QY 300 ILGNIRQAHCNTRAQWNTLQOIATTLREQGKNTIAFNQSGGDEPVEVHSGGGEF 359
 Db 325 ILGDIRQAHCNLSRAKNDTLNKIVIKLREQGKNTIVFKHSSGGDEPVEVHSGGGEF 384
 QY 360 FYCNSQTLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVLPK 419
 Db 385 FYCNSQTLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVLPK 438
 QY 420 GLIRCSSNITGLLLTRDGGENOTTEIFRPGGGDMRDNRSELYKYKVVKIPLGVAPTK 479
 Db 439 GQIRCSSNITGLLLTRDGGEPDNKTEVFRPGGGDMRDNRSELYKYKVVKIPLGVAPTK 498

QY 480 AKRRVVQREKRAVGMGAMPLGFLGAAGSTMGATSMALTVAQRLLSGIVQQNNLLRAI 539
 Db 499 AKRRVVQREKRAVG-IGAVFLGFLGAAGSTMGAAATLTVAQRLLSGIVQQNNLLRAI 557
 QY 540 KAOQHLLQTLVWGIKOLOARILAVERYLKDQOLGFWGCGSKLICCTAVPWNASWNSKTL 599
 Db 558 EAQOHLQTLVWGIKOLOARILAVERYLKDQOLGFWGCGSKLICCTAVPWNASWNSKSL 617
 QY 600 DQIWNNTWMEWDREIDNTHLYIYTBESQOQKNOQELQLDKWASL 649
 Db 618 NKIWNMTWIEWDREINNTYSIYTBESQOQKNOQELQLDKWASL 667

RESULT 7
 US-07-956-483-14
 ; Sequence 14, Application US/07956483
 ; Patent No. 6261799
 ; GENERAL INFORMATION:
 ; APPLICANT: KIENY, Marie-Paule
 ; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
 ; TITLE OF INVENTION: gp160 VARIANT
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/956,483
 ; FILING DATE: 31-DEC-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO 92/19742
 ; FILING DATE: 12-NOV-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 91 05392
 ; FILING DATE: 02-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crane-Feury, Sharon E
 ; REGISTRATION NUMBER: 36,113
 ; REFERENCE/DOCKET NUMBER: 017753-005
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 855 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-07-956-483-14

Query Match 81.1%; Score 2818; DB 3; Length 855;
 Best Local Similarity 81.3%; Pred. No. 1.6e-224;
 Matches 534; Conservative 39; Mismatches 54; Indels 30; Gaps 7;
 QY 3 LWVTYVGVVWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
 Db 33 LWVTYVGVVWKEATTTLCFASDAKAYDTEVHNWATHACVPTNPQEVVLENTNF 92
 QY 63 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWVEQ 122
 Db 93 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTNTTNTS-----SN 148

QY 123 RGK---GEMRNCSENIITTSIRDKVQREYALFYKLDVDPIDDKNTTNTTKYRLINCVTSV 179
Db 149 RGNVEGEMTNCSENIITTSIRSKVQREYALFYKLDVDPID-----NTSTLINCVTSV 201
QY 180 ITQACPKVSEFPIPIHYCTPTGFAALLKCNDRKFKNGTGPCTNVSTVQCTHGRPVVSTQLL 239
Db 202 ITQACPKVSEFPIPIHYCA-RWFALLCNCKFKNGTGPCTNVSTVQCTHGRPVVSTHLL 260
QY 240 LNSLABEEVVRSENFNTNNAKTIIVQLNVSEINCTRPNNHTKRVTLGPRVWYITGE 299
Db 261 LNSLABEEVVRSENFNTNNAKTIIVQLKEAVEINCTRPNNHTKRVTLGPRVWYITGE 320
QY 300 ILGNIRQAHCHNISRAQWNTTQQIATTLREOFNGKTIATFNOSGSDPPIVMHSENCGGEF 359
Db 321 IIGDIRQAHCHNISRAQWNTTQQIATTLREOFNGKTIATFNOSGSDPPIVMHSENCGGEF 380
QY 360 FYCNSTQLFNSAWNT--SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYALP 417
Db 381 FYCNSTQLFNSAWNT--SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYALP 432
QY 418 IKGLIRCSSNITGLLLTRDGG---GGENQTEIFRPGGDMRDNRSELYKYKVVKIEP 472
Db 433 IKGVKCSSNITGLLLTRDGG---GGENQTEIFRPGGDMRDNRSELYKYKVVKIEP 492
QY 473 LGVAPTAKRRVQREKRAVGMGLGAMFLGELGAAGSTWGATSMALTVQAROLLGSGIVQQ 532
Db 493 LGVAPTAKRRVQREKRAVGMGLGAMFLGELGAAGSTWGATSMALTVQAROLLGSGIVQQ 552
QY 533 NNILRAIKAQOHLQLTVWGIKQARILAVERYLKQOQLLGFPGCSGKLICTTAVPWNNA 592
Db 553 NNILRAIKAQOHLQLTVWGIKQARILAVERYLKQOQLLGFPGCSGKLICTTAVPWNNA 612
QY 593 SWSNKTLDQIWNMTWMEWDRIDNYTHLYTLIEESQOQKQOQLLDKQWASL 649
Db 613 SWSNKTLDQIWNMTWMEWDRIDNYTHLYTLIEESQOQKQOQLLDKQWASL 669

RESULT 8

US-08-375-510-2
; Sequence 2, Application US/08375510
; Patent No. 5576421
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo
; TITLE OF INVENTION: HIV ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,510
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/985,949
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 216-309P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848

TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYDROTHERMAL: NO
; ANTI-SENSE: NO
US-08-375-510-2

Query Match 81.0%; Score 2814.5; DB 1; Length 826;
Best Local Similarity 82.9%; Pred. No. 2.9e-224;
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;

QY 3 LWTVTYGVVPVWKEATTTLCASDAKAYDTAEHNVWATHACVTPNPDPQEVVLNVITNF 62
Db 6 LWTVTYGVVPVWKEATTTLCASDAKAYDTAEHNVWATHACVTPNPDPQEVVLNVITNF 65
QY 63 NMKNVVEQHEHDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVVWEQ 122
Db 66 NMKNVVEQHEHDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTSSSGRMINE- 122
123 RGKEMRNCSENIITTSIRDKVQREYALFYKLDVDPIDDKNTTNTTKYRLINCVTSVITQ 182
Db 123 --KGEIKNCSENIITTSIRDKVQREYALFYKLDVDPID-----NTSYRLISCVTSVITQ 173
QY 183 ACQKVSPEPIPIHYCTPTGFAALLKCNDRKFKNGTGPCTNVSTVQCTHGRPVVSTQLLNG 242
Db 174 ACQKVSPEPIPIHYCAPAGFAILLKCNDRKFKNGTGPCTNVSTVQCTHGRPVVSTQLLNG 233
QY 243 SLAEEVVRSENFNTNNAKTIIVQLNVSEINCTRPNNHTKRVTLGPRVWYITGEI 300
Db 234 SLAEEVVRSENFNTNNAKTIIVQLNVSEINCTRPNNHTKRVTLGPRVWYITGEI 293
QY 301 LGNIRQAHCHNISRAQWNTTQQIATTLREOFNGKTIATFNOSGSDPPIVMHSENCGGEF 359
Db 294 --GNRQAHCHNISRAQWNTTQQIATTLREOFNGKTIATFNOSGSDPPIVTHSENCGGEF 352
QY 360 FYCNSTQLFNSAWNTSNGTWSVTRKQKDTG-DIITLPCRIKQIINRWQVVGKAMYALPI 418
Db 353 FYCNSTQLFNSAWNTSNGTWSVTRKQKDTG-DIITLPCRIKQIINRWQVVGKAMYALPI 409
QY 419 KGLIRCSSNITGLLLTRDGG---GGENQTEIFRPGGDMRDNRSELYKYKVVKIEP 478
Db 410 KGLIRCSSNITGLLLTRDGG---GGENQTEIFRPGGDMRDNRSELYKYKVVKIEP 469
QY 479 KAKRRVQREKRAVGMGLGAMFLGELGAAGSTWGATSMALTVQAROLLGSGIVQQNNLLRA 538
Db 470 KAKRRVQREKRAVGMGLGAMFLGELGAAGSTWGATSMALTVQAROLLGSGIVQQNNLLRA 528
QY 539 IKAQOHLQLTVWGIKQARILAVERYLKQOQLLGFPGCSGKLICTTAVPWNNAWSNKT 598
Db 529 IKAQOHLQLTVWGIKQARILAVERYLKQOQLLGFPGCSGKLICTTAVPWNNAWSNKT 588
QY 599 LDOJWNMTWMEWDRIDNYTHLYTLIEESQOQKQOQLLDKQWASL 649
Db 589 LDOJWNMTWMEWDRIDNYTHLYTLIEESQOQKQOQLLDKQWASL 639

RESULT 9

US-08-487-657-2
; Sequence 2, Application US/08487657
; Patent No. 5834267
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo
; TITLE OF INVENTION: HIV ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch

STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,657
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/375,510
FILING DATE: 18-JAN-1995
APPLICATION NUMBER: US/07/985,949
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 216-309P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-487-657-2

Query Match 81.0%; Score 2814.5; DB 2; Length 826;
Best Local Similarity 82.9%; Pred. No. 2.9e-224;
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;
QY 3 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQVVLNVNTENF 62
DB 6 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQVVLNVNTENF 65
QY 63 NMKNMNVQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVVEQ 122
DB 66 NMKNMNVQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTSSSGRIME- 122
QY 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNSTVITQ 182
DB 123 --KGEIKNCSFNISTIRDKVQREYAFYKLDIVPID-----NTSYRLISCNSTVITQ 173
QY 183 ACPKVSFEPIPIHYCTPTGFPALLKCNDFKNGTGPCTNVSTVQCTHGRIPVWSTQLLNG 242
DB 174 ACPKVSFEPIPIHYCAPAGFAILLKCNKNTFNGTGPCTNVSTVQCTHGRIPVWSTQLLNG 233
QY 243 SLAEEVWIRSENFTNNAKTIIVQLNVSVVEINCTRNPNHTRK--RVTLGPRVWYTTGEI 300
DB 234 SLAEDVWIRSANFTDNAKTIIVQLNVSVEINCTRNPNHTRKSIIRQSGGDEPIVTHSFNCGGEF 293
QY 301 LGNIROAHNISRAQWNTLQIATTLREQFG-NKTIAFNQSOGGDEPIVWHSFNCGGEF 359
DB 294 -GNMROAHNISRAKWNTLQIASKLREQFGNKTIIIFKQSSGDEPIVTHSFNCGGEF 352
QY 360 FYCNSSTLFNSAMVNTSGTWSVTRKQDGTG-DIITLPCRKQIINRWQVGVKAMVALPI 418
DB 353 FYCNSSTLFNSW---FNSTWSTEGSNNTGSDTIILPCRKQIINRWQVGVKAMVALPI 409
QY 419 KGLIRCSNITGLLLTRDGGNGTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478
DB 410 SGQIRCSNITGLLLTRDGGNNNGSIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 469

QY 479 KAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 470 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGCTSMTLTVOARQLLSGIVQOQNLLRA 528
QY 539 IKAQOHLQLTWGIGIKQLOARILAVERYLKQOQLLGFWGCCKLICTTAVPWNASWSNKT 598
DB 529 IEAQOHLQLTWGIGIKQLOARILAVERYLKQOQLLGFWGCCKLICTTAVPWNASWSNKS 588
QY 599 LDQIWNMTWMEWDREIDNTHLYTLIESQOQEKNOQELLQLDKWSL 649
DB 589 LEQIWNMTWMEWDREINNTSLIHSILIESQOQEKNOELLELDKWSL 639

RESULT 10

US-09-309-572-23
Sequence 23, Application US/09309572
Patent No. 6440730
GENERAL INFORMATION:
APPLICANT: Heinrich-Pette-Institut
TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
FILE REFERENCE: P50489
CURRENT APPLICATION NUMBER: US/09/309,572
CURRENT FILING DATE: 1999-05-11
EARLIER APPLICATION NUMBER: DE 198 56 463
EARLIER FILING DATE: 1998-11-26
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 854
TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
FEATURE:
OTHER INFORMATION: envelope polyprotein
US-09-309-572-23

Query Match 81.0%; Score 2814.5; DB 4; Length 854;
Best Local Similarity 82.9%; Pred. No. 3.1e-224;
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;
QY 3 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQVVLNVNTENF 62
DB 34 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQVVLNVNTENF 93
QY 63 NMKNMNVQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVVEQ 122
DB 94 NMKNMNVQMHEDIISLWQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150
QY 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNSTVITQ 182
DB 151 --KGEIKNCSFNISTIRDKVQREYAFYKLDIVPID-----NTSYRLISCNSTVITQ 201
QY 183 ACPKVSFEPIPIHYCTPTGFPALLKCNDFKNGTGPCTNVSTVQCTHGRIPVWSTQLLNG 242
DB 202 ACPKVSFEPIPIHYCAPAGFAILLKCNKNTFNGTGPCTNVSTVQCTHGRIPVWSTQLLNG 261
QY 243 SLAEEVWIRSENFTNNAKTIIVQLNVSVVEINCTRNPNHTRK--RVTLGPRVWYTTGEI 300
DB 262 SLAEDVWIRSANFTDNAKTIIVQLNVSVEINCTRNPNHTRKSIIRQSGGDEPIVTHSFNCGGEF 321
QY 301 LGNIROAHNISRAQWNTLQIATTLREQFG-NKTIAFNQSOGGDEPIVWHSFNCGGEF 359
DB 322 -GNMROAHNISRAKWNTLQIASKLREQFGNKTIIIFKQSSGDEPIVTHSFNCGGEF 380
QY 360 FYCNSSTLFNSAMVNTSGTWSVTRKQDGTG-DIITLPCRKQIINRWQVGVKAMVALPI 418
DB 381 FYCNSSTLFNSW---FNSTWSTEGSNNTGSDTIILPCRKQIINRWQVGVKAMVALPI 437
QY 419 KGLIRCSNITGLLLTRDGGNGTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478
DB 438 SGQIRCSNITGLLLTRDGGNNNGSIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 497
QY 479 KAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538

Db 498 KAKRRVQEKRAVG-IGALFLGFLGAAGSTMGCTSMITVQARQLSDIVQOQNLLRA 556
QY 539 IKAQOHLQLTVWGKIQLOARILAVERYLKQOQLLGFWSGSKLICTTAVPWNASWNSKT 598
Db 557 IEAQOHLQLTVWGKIQLOARILAVERYLKQOQLLGFWSGSKLICTTAVPWNASWNSKS 616
QY 599 LQIWNMTWMDREIDNTHYIYTLIBESQOQEKNOQELLQLDKWASL 649
Db 617 LEQIWNMTWMDREIDNTHYIYTLIBESQOQEKNOQELLQLDKWASL 667

RESULT 11
US-08-788-815-7
; Sequence 7, Application US/08788815
; Patent No. 5846546
; GENERAL INFORMATION:
; APPLICANT: St. Jude Children's Research Hospital
; APPLICANT: 332 No. 5846546th Lauderdale
; APPLICANT: PO Box 318
; APPLICANT: Memphis, TN 38101-0318
; APPLICANT: United States of America
; APPLICANT: Coleclough, Christopher
; APPLICANT: Owens, Randall J.
; APPLICANT: Slobod, Karen
; TITLE OF INVENTION: PREPARATION AND USE OF VIRAL VECTORS FOR
; TITLE OF INVENTION: MIXED ENVELOPE PROTEIN VACCINES AGAINST HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: KLAUBER & JACKSON
; STREET: 411 HACKENSACK AVENUE
; CITY: HACKENSACK
; STATE: NJ
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,815
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehlner
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: 1340-1-011CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5500
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 880 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: protein
; US-08-788-815-7

Query Match 81.0%; Score 2814; DB 2; Length 880;
Best Local Similarity 82.1%; Pred. No. 3.5e-224;
Matches 540; Conservative 39; Mismatches 61; Indels 18; Gaps 9;
QY 3 LWTVYGVVPVWKEATITLFCASAKAYDTEAHNVWATHACVPNPNPQEVVLENTVF 62
Db 43 LWTVYGVVPVWKEATITLFCASAKAYDTEAHNVWATHACVPNPNPQEVVLENTVF 102
QY 63 NMWKNMVQOMHEDIISLWDSQKPCVKLTCLVLTNCTDL--NNNTNTTETLSIIWV 120
Db 103 NMWKNMVQOMHEDIISLWDSQKPCVKLTCLVLTNCTDL--NNNTNTTETLSIIWV 158

QY 121 EQR--CKGEMRNCNFIITSIKDVQREYALFYKLDVEPID--DNKNTNTNTKYLINCN 176
Db 159 GRNIMEEGIKNCNFIITSIKDVQREYALFYKLDIIPIDKNDNDNTSYKFTITSCN 218
QY 177 TSVITQACPKVSEPIPIHYCTPTGALLKNDKFNKGTGPTCTNVSTVQCTHGIRPVST 236
Db 219 TSVITQACPKVSEPIPIHYCAPAGFAILKCNKTFNGTGTCTNVSTVQCTHGIRPVST 278
QY 237 QLLNGSLAEERVIRSENFTNNAKTIIVQLNVSEINCTRENHTRK--RVTLGSGRVW 294
Db 279 QLLNGSLAEERVIRSANFTDNAKTIIVQLNQSVSEINCTRENHTRKIRIQRGFRAF 338
QY 295 YTTGELGNIRQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNOSGGDPEIWHSPF 353
Db 339 VTIGKILGNWQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNOSGGDPEIWHSPF 398
QY 354 NCGGEFFYCNSTQLFNSAWNVTSTGWSVTRKQKDTG-DIITLPCRKQIINRWQVVGKA 412
Db 399 NCGGEFFYCNSTQLFNSAWNVTSTGWSVTRKQKDTG-DIITLPCRKQIINRWQVVGKA 455
QY 413 MYALPIKGLIRGSSNITGLILTRDGG-GENQTTETIFRPGGDMRDNWSELYKYKVVKIE 471
Db 456 MYALPIKGLIRGSSNITGLILTRDGG-GENQTTETIFRPGGDMRDNWSELYKYKVVKIE 515
QY 472 PLGVAPTAKRRVVQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQO 531
Db 516 PLGVAPTAKRRVVQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQO 575
QY 532 QNNLLRAIKAOQHLLQLTVWGKIQLOARILAVERYLKQOQLLGFWSGSKLICTTAVPWN 591
Db 576 QNNLLRAIKAOQHLLQLTVWGKIQLOARILAVERYLKQOQLLGFWSGSKLICTTAVPWN 635
QY 592 ASWSNKTLDQIWNMTWMDREIDNTHYIYTLIBESQOQEKNOQELLQLDKWASL 649
Db 636 ASWSNKTLDQIWNMTWMDREIDNTHYIYTLIBESQOQEKNOQELLQLDKWASL 693

RESULT 12
US-09-157-963-7
; Sequence 7, Application US/09157963
; Patent No. 6086891
; GENERAL INFORMATION:
; APPLICANT: St. Jude Children's Research Hospital
; APPLICANT: 332 No. 6086891th Lauderdale
; APPLICANT: PO Box 318
; APPLICANT: Memphis, TN 38101-0318
; APPLICANT: United States of America
; APPLICANT: Coleclough, Christopher
; APPLICANT: Owens, Randall J.
; APPLICANT: Slobod, Karen
; TITLE OF INVENTION: PREPARATION AND USE OF VIRAL VECTORS FOR
; TITLE OF INVENTION: MIXED ENVELOPE PROTEIN VACCINES AGAINST HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: KLAUBER & JACKSON
; STREET: 411 HACKENSACK AVENUE
; CITY: HACKENSACK
; STATE: NJ
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,963
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehlner
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: 1340-1-011CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5500
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 880 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: protein
; US-09-157-963-7

```
; REFERENCE/DOCKET NUMBER: 13401011 CP DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 880 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: protein
; US-09-157-963-7

Query Match      81.0%; Score 2814; DB 3; Length 880;
Best Local Similarity 82.6%; Pred. No. 3.5e-224;
Matches 540; Conservative 39; Mismatches 61; Indels 18; Gaps 9;

QY 3 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTNPNPOEVVLNVTENF 62
Db 43 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTDNPQEVVLNVTENF 102
QY 63 NWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDL--NTNNTNTNTELSIIIVW 120
Db 103 NWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTSNNVTSS---W 158
QY 121 EQR--CKGEMRNCSENIITSIKDKVQREYALFYKLDVEPID--DNKNTNTNXYRLINC 176
Db 159 GRNIMEEGSIKSNFNIISIRKGVQKEYAFFYKLDIIPIDKGNDSNDTSYKFTLTSN 218
QY 177 TSVITQACPKVSEPIPIHYCTPTGTPGALLKCNCKFKNGTGPCNVTSTVQCTHGIRPVST 236
Db 219 TSVITQACPKVSEPIPIHYCAPAGAILKCNKNTFNGTGPCNVTSTVQCTHGIRPVST 278
QY 237 QLLNGSLAEEVWIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPRVW 294
Db 279 QLLNGSLAEEVWIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKSIIRIQGFGRF 338
QY 295 YTTGETILNIRQAHNCSRAQNNNTLQOATTLREQFG--NKTIAPNQSOGDPEIIVHSP 353
Db 339 VTIGKILGNRQAHNCSRAQNNNTLQOATTLREQFG--NKTIAPNQSOGDPEIIVHSP 398
QY 354 NCGGEFFYCNSTQLFNSANVTISNGTWSVTRKQKDTG--DIITLPCRILKOLINRWQVVGKA 412
Db 399 NCGGEFFYCNSTQLFNSW---FNSWTSTKGSNTEGSDTILPCRILKOLINRWQVVGKA 455
QY 413 MYALPIKGIHRCSSNTIGLLTRDGG--GENQTTPEIFRPGGDMRDNRSELYKYKVKIE 471
Db 456 MYAPISGQIRCSSNTIGLLTRDGGANENNESEIFRPGGDMRDNRSELYKYKVKIE 515
QY 472 PLGVAPTKAKRVQREKAVGMLGAMFLGFLGAAGSTWGATSMALTVOARQLLSIVQ 531
Db 516 PLGVAPTKAKRVQREKAVGMLGAMFLGFLGAAGSTWGATSMALTVOARQLLSIVQ 575
QY 532 QNNLLRAIKAQHLLQTLVWGIKQARILAVERYLKDOQLLGFWCSCGKLICTTAVPNW 591
Db 576 QNNLLRAIKAQHLLQTLVWGIKQARILAVERYLKDOQLLGFWCSCGKLICTTAVPNW 635
QY 592 ASWSNKLDOIWNMTWMDREIDNYTHLIYTLIEESQOQEKNOQELLDKNASL 649
Db 636 ASWSNKLDOIWNMTWMDREIDNYTHLIYTLIEESQOQEKNOQELLDKNASL 693

RESULT 13
US-07-916-098A-2
; Sequence 2, Application US/07916098A
; Patent No. 5871732
; GENERAL INFORMATION:
; APPLICANT: BURKLY, LINDA C.
; APPLICANT: CHISHOLM, PATRICIA L.
; APPLICANT: THOMAS, DAVID W.
; APPLICANT: ROSA, MARGARET D.
; APPLICANT: ROSA, JOSEPH J.
; TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
```

```
; TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
; STREET: 10 SOUTH WACKER DRIVE
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/916,098A
; FILING DATE: July 24, 1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08843
; FILING DATE: NO. 5871732ember 27, 1991
; CLASSIFICATION: 424
; APPLICATION NUMBER: 07/618,542
; FILING DATE: NO. 5871732ember 27, 1990
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHN J. MC DONNELL
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,310-G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 715-1000
; TELEFAX: (312) 715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-916-098A-2

Query Match      80.9%; Score 2809.5; DB 2; Length 856;
Best Local Similarity 82.6%; Pred. No. 7.9e-224;
Matches 538; Conservative 38; Mismatches 56; Indels 19; Gaps 9;

QY 3 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTNPNPOEVVLNVTENF 62
Db 34 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTDNPQEVVLNVTENF 93
QY 63 NWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTNTELSIIIVW 122
Db 94 NWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDTNTNSSSGRIME- 150
QY 123 RKGEMRNCSENIITSIKDKVQREYALFYKLDVEPIDDNKNTNTNTELSIIIVW 182
Db 151 --KGEIKNCSENIITSIKDKVQREYALFYKLDIIPIDN----DFTSYKLTSCNVTIQ 203
QY 183 ACPKVSFEPIPIHYCTPTGTFALLKCNCKFKNGTGPCNVTSTVQCTHGIRPVVSTOLLNG 242
Db 204 ACPKVSFEPIPIHYCAPAGAILKCNKNTFNGTGPCNVTSTVQCTHGIRPVVSTOLLNG 263
QY 243 SLABEEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTL--GPGRVVYTTGEI 300
Db 264 SLABEEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRIQRGPGRAFVTIGKI 323
QY 301 LGNIRQAHNCSRAQNNNTLQOATTLREQFG--NKTIAPNQSOGDPEIIVHSP 359
Db 324 --GNMRQAHNCSRAQNNNTLQOATTLREQFG--NKTIAPNQSOGDPEIIVHSP 382
QY 360 FYCNSDTQLFNSANVTISNGTWSVTRKQKDTG--DIITLPCRILKOLINRWQVVGKA 418
Db 383 FYCNSDTQLFNSW---FNSWTSTEGSNTEGSDTILPCRILKOLINRWQVVGKA 439
```


QY 419 KGLIRCSNITGLLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 478
DB 440 SQIIRCSNITGLLLTRDGGGNSNESEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 499
QY 479 KAKRRVQREKRAVGMGLGAMFLGFGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 500 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGAASMTLTVOARQLLSGIVQOQNLLRA 558
QY 539 IKAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGWCSGKLICTTAVPNWASWSNKT 598
DB 559 IEAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGWCSGKLICTTAVPNWASWSNKS 618
QY 599 LDOIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLOLDKWASL 649
DB 619 LEQIWNHTTWMEWDREIDNTHLYTLIEESQOQKNOQELLOLDKWASL 669

RESULT 14
US-08-448-603A-30
; Sequence 30, Application US/08448603A
; Patent No. 5864027
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,603A
; FILING DATE: 07-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/072,833
; FILING DATE: 07-JUN-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Halliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-448-603A-30

Query Match 80.9%; Score 2809.5; DB 2; Length 857;
Best Local Similarity 81.3%; Pred No. 7.9e-224;
Matches 529; Conservative 45; Mismatches 60; Indels 17; Gaps 6;

QY 3 LWVTYYGVPMKEATTTTLCFASDAXDYDEAHNWMATHACVPTNPNFQEVVLENTVF 62
DB 33 LWVTYYGVPMKETTTLTLCFASDAXDYDETHNWMATHACVPTDNPQEVVLENTVF 92
QY 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVNHEQ 122
DB 93 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD--AGNTTNTNSSS----REK 146

QY 123 RKGEMRNCSENIITTSIRDKVQREYALFKLDVEPIDDN--KNTTNTKYLLJNCNTSVI 180
DB 147 LEKGEIKNCSENIITTSVEDKVKQETALFNKLDIVFDODDDRNSTNRSTNYRLISCNTSVI 206
QY 181 TOACPQKVSFEPIPIHYCTPTGTFALLKCNDDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLL 240
DB 207 TOACPQKVSFEPIPIHFCTPAGFALLKCNKNTFNGSGPKNVSTVQCTHGIRPVVSTQLLL 266
QY 241 NGSLAEERVWIRSENFTNNAKTIIVQLNVSYEINCTRENNTNRKVTIPGRVWVYTTGEI 300
DB 267 NGSLAEGERVWIRSENFTNNAKTIIVQLTEPKINCTRENNTNRKIPGPGRAFATGDI 326
QY 301 LGNIRQAHCNISRAQWNNLTQOIIATTLREQFNGKNTIAFNQSSGGDPEIWMHSFNCGGEFF 360
DB 327 LGNIRQAHCNLSRTDNNNTLQIIVKELREQFNGKNTIIEHSSGGDPEIWMHSFNCRGEFF 386
QY 361 YCNSTQLPNSAWNV--SNGTWSVTRKOKDGTGDIITLPCRKIKOILNRQVVGKAMVALPI 418
DB 387 YCNTTQLFDSTWDNTKVSNGT-----STEENSTITLPCRKIKOIVNMWQEVGKAMYAPPI 440
QY 419 KGLIRCSNITGLLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 478
DB 441 RGQIRCSNITGLLLTRDGGGNSNNEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 500
QY 479 KAKRRVQREKRAVGMGLGAMFLGFGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 501 KAKRRVQREKRAVG-IGAVFLGFLGAAGSTMGAASITLTVOARQLLSGIVQOQNLLRA 559
QY 539 IKAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGWCSGKLICTTAVPNWASWSNKT 598
DB 560 IEAQOHLIQLIIVWGIKQIARILAVERYLKQOQLLGFPGWCSGKLICTTAVPNWASWSNKS 619
QY 599 LDOIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLOLDKWASL 649
DB 620 LDKIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLOLDKWASL 670

RESULT 15
US-09-134-075-30
; Sequence 30, Application US/09134075
; Patent No. 6042836
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; APPLICANT: Nakamura, Gerald R.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP
; STREET: 3 Embarcadero Center
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,075
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/448,603
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halliday, Emily
; REGISTRATION NUMBER: 38903
; REFERENCE/DOCKET NUMBER: 14918-704
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-393-2000
; TELEFAX: 415-393-2286
; TELEX:

; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-134-075-30

```
Query Match      80.9%; Score 2809.5; DB 3; Length 857;
Best Local Similarity 81.3%; Pred. No. 7.9e-224;
Matches 529; Conservative 45; Mismatches 60; Indels 17; Gaps 6;

QY      3  LNVTVYGVGVWKEATTTLFCASDAKAYDTEAHNWMATHACVPTNPPOEVVLENTENF 62
Db      33 LNVTVYGVGVWKEATTTLFCASDAKAYDTEAHNWMATHACVPTNPPOEVVLENTENF 92

QY      63 NNMKNMVEQMEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTSLIIVWVEQ 122
Db      93 NNMKNMVEQMEHDIISLWDQSLKPCVKLTPLCVTLNCTD--AGNTTNTNSS----REK 146

QY      123 RKGEMRNCSPNITSIRDKVOREVALFYKLDVEPIDDK--KNTTNNTKYRLINCNTSVI 180
Db      147 LEKGIKKCSFNITTSVRDKMKETALEFNKLDIVPIDDDDRNSTRNSTNYRLISCNTSVI 206

QY      181 TCACPKVSEFPIPIHYCTPTGFALLKCNCKXKNGTGTNVSTVCTGTHGIRPVVSTQLLL 240
Db      207 TCACPKVSEFPIPIHYCTPTGFALLKCNCKXKNGTGTNVSTVCTGTHGIRPVVSTQLLL 266

QY      241 NGSLEAEVWIRSENFNTNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPGRVWYTTGEI 300
Db      267 NGSLEAEVWIRSENFNTNAKTIIVQLTEPVKINCTRPNNHTRKRSIPGPGRAFVATGDI 326

QY      301 LGNTRQAHNCSRAQWNTLQIATILREQFGNKTIAFNQSSGGDPEIVMHSFNCGGEPF 360
Db      327 LGNTRQAHNCSRAQWNTLQIATILREQFGNKTIIFNHSSGGDPEIVMHSFNCRGEFF 386

QY      361 YCNSTQLFNSAWNVT--SNGTWSVTRKOKDGTGDIITLPCRICKIINRMQVGVKAMYALPI 418
Db      387 YCNSTQLFNSAWNVT--SNGTWSVTRKOKDGTGDIITLPCRICKIINRMQVGVKAMYALPI 440

QY      419 KGLIRCSSNITGLLITRDGGNGENQTTEIPRPGGDMRDNRSELYKYVVKIEPLGVAPT 478
Db      441 RQGIKRCSSNITGLLITRDGGNSNMNETPRPGGDMRDNRSELYKYVVKIEPLGVAPT 500

QY      479 KAKRVVQREKAEVGMGLGAMFLGFLGAAGSTWGAATSMALTVOARQLLSGIVQQNNLLRA 538
Db      501 KAKRVVQREKRAVG--IGAVFLGFLGAAGSTWGAASITLTVQARLLLSGIVQQNNLLRA 559

QY      539 IKAQQHLLQLTWVGIKQLQARILAVERYLKDQQLLGFWGCCKGLICTTAVPWNASWSNKT 598
Db      560 IEAQHLLQLIIVWGIKQLQARILAVERYLRDQQLLGIWGCCKGLICTTSVPWNASWSNKS 619

QY      599 LDQIWNNTMWEWDREIDNYTHLIYTLIEESQNOQEKNOQELLQDLQKWSL 649
Db      620 LDKIWNNTMWEWDREIENYTSLIYTLIEESQNOQEKNEQDLLELDQWASL 670
```

Search completed: January 13, 2004, 09:12:33
Job time : 24 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:10:37 ; Search time 38 Seconds

(without alignments)
3441.574 Million cell updates/sec

Title: US-09-938-406-1_COPY_33_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVPMKEATT.....QNQKKNQQLQLDKWASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3474	100.0	868	10	US-09-938-406-1
2	2883.5	83.0	842	12	US-10-190-435-2
3	2883.5	83.0	842	12	US-10-241-009-2
4	2883.5	83.0	842	12	US-10-190-434B-2
5	2883.5	83.0	842	12	US-10-190-305A-2
6	2883.5	83.0	847	10	US-09-476-242-2
7	2827.5	81.4	643	15	US-10-032-162-13
8	2809.5	80.9	856	10	US-09-476-242-1
9	2795.5	80.5	856	15	US-10-196-515-11
10	2795.5	80.5	861	15	US-10-026-741-103
11	2757	79.4	726	15	US-10-196-515-3
12	2755	79.3	759	15	US-10-196-515-12
13	2728.5	78.5	619	10	US-09-891-609-4
14	2728.5	78.5	646	10	US-09-891-609-2
15	2720.5	78.3	625	15	US-10-032-162-17

16	2585	74.4	860	12	US-10-190-435-6	Sequence 6, Appli
17	2585	74.4	860	12	US-10-241-009-6	Sequence 6, Appli
18	2585	74.4	860	12	US-10-190-434B-6	Sequence 6, Appli
19	2585	74.4	860	12	US-10-190-305A-6	Sequence 6, Appli
20	2577.5	74.2	853	12	US-10-369-294-13	Sequence 13, Appl
21	2566	73.9	855	12	US-10-190-435-144	Sequence 144, App
22	2563.5	73.8	858	12	US-10-190-435-150	Sequence 150, App
23	2557.5	73.6	579	15	US-10-032-162-15	Sequence 15, Appl
24	2553	73.5	870	12	US-10-190-435-127	Sequence 127, App
25	2552.5	73.5	867	12	US-10-190-435-3	Sequence 3, Appli
26	2552.5	73.5	867	12	US-10-241-009-3	Sequence 3, Appli
27	2552.5	73.5	867	12	US-10-190-434B-3	Sequence 3, Appli
28	2552.5	73.5	867	12	US-10-190-305A-3	Sequence 3, Appli
29	2552.5	73.5	867	12	US-10-190-435-126	Sequence 126, App
30	2551	73.4	855	12	US-10-369-294-11	Sequence 11, Appl
31	2551	73.4	855	12	US-10-190-435-129	Sequence 12, Appl
32	2550	73.4	845	12	US-10-190-435-130	Sequence 129, App
33	2550	73.4	845	12	US-10-190-435-138	Sequence 130, App
34	2548.5	73.4	869	12	US-10-241-009-4	Sequence 4, Appli
35	2548.5	73.4	869	12	US-10-190-434B-4	Sequence 4, Appli
36	2548.5	73.4	869	12	US-10-190-305A-4	Sequence 4, Appli
37	2548.5	73.4	861	12	US-10-190-435-139	Sequence 139, App
38	2542.5	73.2	865	12	US-10-190-435-140	Sequence 140, App
39	2541	73.1	857	12	US-10-190-435-142	Sequence 142, App
40	2531	72.9	862	12	US-10-190-435-141	Sequence 141, App
41	2529	72.8	862	12	US-10-190-435-143	Sequence 143, App
42	2528.5	72.8	844	10	US-09-991-258-19	Sequence 19, Appl
43	2524.5	72.7	862	12	US-10-190-435-141	Sequence 141, App
44	2523	72.6	845	12	US-10-190-435-143	Sequence 143, App
45	2519	72.5	803	12	US-10-190-435-134	Sequence 134, App

ALIGNMENTS

RESULT 1

US-09-938-406-1 ; Sequence 1, Application US/09938406
; Patent No. US20020155120A1
; GENERAL INFORMATION:
; APPLICANT: Lowell, George
; APPLICANT: Vancott, Thomas
; APPLICANT: Birx, Deborah
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR
; TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY
; FILE REFERENCE: 40646-20002.10
; CURRENT APPLICATION NUMBER: US/09/938,406
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/214,701
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: PCT/US 97/12253
; PRIOR FILING DATE: 1997-07-10
; PRIOR APPLICATION NUMBER: US 60/021,687
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Virus HIV-1
US-09-938-406-1

Query Match 100.0%; Score 3474; DB 10; Length 868;
Best Local Similarity 100.0%; Pred. No. 5.4e-306;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANLWTVYGVVPMKEATTTLFCASDAKAYTEAHNWTACVTPNPNQEVLENVTE 60
DB 33 ANLWTVYGVVPMKEATTTLFCASDAKAYTEAHNWTACVTPNPNQEVLENVTE 92
QY 61 NFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVW 120
DB 93 NFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVW 152

QY 121 EORKGEMRNCNFNITTSIRDKVOREVALFYKLDVEPIDDKNTNTNTKYRLINCNTSVI 180
 DB 153 EORKGEMRNCNFNITTSIRDKVOREVALFYKLDVEPIDDKNTNTNTKYRLINCNTSVI 212
 QY 181 TOACPKVSFEPIPIHYCTPTGFPALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 240
 DB 213 TOACPKVSFEPIPIHYCTPTGFPALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 272
 QY 241 NGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEI 300
 DB 273 NGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEI 332
 QY 301 LGNIRQAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNOSGGDPEIWMHSFNCGGEFF 360
 DB 333 LGNIRQAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNOSGGDPEIWMHSFNCGGEFF 392
 QY 361 YCNSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMYALPIKG 420
 DB 393 YCNSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMYALPIKG 452
 QY 421 LIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKA 480
 DB 453 LIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKA 512
 QY 481 KRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOONLLRAIK 540
 DB 513 KRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOONLLRAIK 572
 QY 541 AQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGWCGSKLICCTTAVPWNASWNTLID 600
 DB 573 AQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGWCGSKLICCTTAVPWNASWNTLID 632
 QY 601 QIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWASL 649
 DB 633 QIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWASL 681

RESULT 2

US-10-190-435-2

; Sequence 2, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGERDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; CURRENT FILING DATE: 2002-12-30

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SFI62

US-10-190-435-2

Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY

3 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNVWATHACVPTNPNQEVLENVTENF 62

DB

28 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNVWATHACVPTNPNQEVLENVTENF 87

QY

63 NWKKNVMEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTTNTTSLIIVWQ 122

DB

88 NWKKNVMEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTTNTTSLIIVWQ 141

DB 88 NWKKNVMEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTTNTTSLIIVWQ 141
 QY 123 RGKEMRNCNFNITTSIRDKVOREVALFYKLDVEPIDDKNTNTNTKYRLINCNTSVI 182
 DB 142 MDREIKNCSFKVTTTSIRNKWQREYALFYKLDVVDPIDN-----DNTSYKLINCNTSVITQ 196
 QY 183 ACPKVSFEPIPIHYCTPTGFPALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 242
 DB 197 ACPKVSFEPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLL 256
 QY 243 SLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEI 302
 DB 257 SLAEVEVIRSENFNTNNAKTIIVQLKESVEINCTRENHTRKRVTLGPGRVWYTTGEI 316
 QY 303 NROAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNOSGGDPEIWMHSFNCGGEFF 362
 DB 317 DROAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNOSGGDPEIWMHSFNCGGEFF 376
 QY 363 NSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMYALP 417
 DB 377 NSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMYALP 423
 QY 418 IKGLIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGV 476
 DB 424 IRGQIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGV 483
 QY 477 PTKAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOONLL 536
 DB 484 PTKAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOONLL 542
 QY 537 RAIKAQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGWCGSKLICCTTAVPWNASW 596
 DB 543 RAIKAQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGWCGSKLICCTTAVPWNASW 602
 QY 597 KTLDOIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWASL 649
 DB 603 KSLDOIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWASL 655

RESULT 3

US-10-241-009-2

; Sequence 2, Application US/10241009

; Publication No. US20030170614A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGERDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

; FILE REFERENCE: 2300-1621.21

; CURRENT APPLICATION NUMBER: US/10/241,009

; CURRENT FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SFI62

US-10-241-009-2

Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY

3 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNVWATHACVPTNPNQEVLENVTENF 62

DB

28 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNVWATHACVPTNPNQEVLENVTENF 87

QY

63 NWKKNVMEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTTNTTSLIIVWQ 122

DB

88 NWKKNVMEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTTNTTSLIIVWQ 141

QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVERIDDKNNTNTNTKYBLINCNTSVITQ 182
Db 142 MRGEIKNCSEFKVTTTSIRNMQKEYALFYKLDVVPIDN-----DNTSYKLINCNTSVITQ 196
QY 183 ACPKVSFEPIPHYCTPTGTFALLKCKNDKFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 197 ACPKVSFEPIPHYCAPAGFAILKCKNDKFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 256
QY 243 SLAEVWIRSENFTNAKTIIVQLNVSEINCTRPNNHTRKRVLTGPGKRVWYTTGEILG 302
Db 257 SLAEGWIRSENFTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPGRFVATGDIIG 316
QY 303 NTRQACHNISRAOWNNTLOQIATTLREQCNKTIAPNQSSGGDPEIWMHSFNCGGFFYC 362
Db 317 DTRQACHNISGKWNNTLKOIVTKLQAFGNKTIIVFKQSSGGDPEIWMHSFNCGGFFYC 376
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAMYALP 417
Db 377 NSTQLFNSWNTTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYALP 423
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 476
Db 424 IRGQIRCSNITGLLLTRDGGGEISNTTIFRPGGDMRDNWRSELYKYKVKIEPLGVA 483
QY 477 PTKARVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNILL 536
Db 484 PTKARVVQREKRAV-TLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQOQNILL 542
QY 537 RAIKAQHLQLTQVWGIKQIARILAVERYLKDQQLLGFWGCGSKLICTTAVPWNASWN 596
Db 543 RAIKAQHLQLTQVWGIKQIARILAVERYLKDQQLLGFWGCGSKLICTTAVPWNASWN 602
QY 597 KTLDOIWNNTWMEWDREIDNTYTHLIYTBESQOQEKNOQELLDKXWASL 649
Db 603 KSLDOIWNNTWMEWDREIDNTYTHLIYTBESQOQEKNOQELLDKXWASL 655

RESULT 4

US-10-190-434B-2
; Sequence 2, Application US/10190434B
; Publication No. US20030194800A1

GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

; FILE REFERENCE: 2300-1621-20

; CURRENT APPLICATION NUMBER: US/10/190,434B

; CURRENT FILING DATE: 2002-07-05

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-434B-2

Query Match 83.0%; Score 2883.5; DB 12; Length 842;
Best Local Similarity 83.3%; Pred. No. 2e-252;
Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY 3 LWTVYVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 28 LWTVYVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 87
QY 63 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTETLSIIIVWQ 122
Db 88 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 141

QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVERIDDKNNTNTNTKYBLINCNTSVITQ 182
Db 142 MRGEIKNCSEFKVTTTSIRNMQKEYALFYKLDVVPIDN-----DNTSYKLINCNTSVITQ 196
QY 183 ACPKVSFEPIPHYCTPTGTFALLKCKNDKFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 197 ACPKVSFEPIPHYCAPAGFAILKCKNDKFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 256
QY 243 SLAEVWIRSENFTNAKTIIVQLNVSEINCTRPNNHTRKRVLTGPGKRVWYTTGEILG 302
Db 257 SLAEGWIRSENFTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPGRFVATGDIIG 316
QY 303 NTRQACHNISRAOWNNTLOQIATTLREQCNKTIAPNQSSGGDPEIWMHSFNCGGFFYC 362
Db 317 DTRQACHNISGKWNNTLKOIVTKLQAFGNKTIIVFKQSSGGDPEIWMHSFNCGGFFYC 376
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAMYALP 417
Db 377 NSTQLFNSWNTTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYALP 423
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 476
Db 424 IRGQIRCSNITGLLLTRDGGGEISNTTIFRPGGDMRDNWRSELYKYKVKIEPLGVA 483
QY 477 PTKARVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNILL 536
Db 484 PTKARVVQREKRAV-TLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQOQNILL 542
QY 537 RAIKAQHLQLTQVWGIKQIARILAVERYLKDQQLLGFWGCGSKLICTTAVPWNASWN 596
Db 543 RAIKAQHLQLTQVWGIKQIARILAVERYLKDQQLLGFWGCGSKLICTTAVPWNASWN 602
QY 597 KTLDOIWNNTWMEWDREIDNTYTHLIYTBESQOQEKNOQELLDKXWASL 649
Db 603 KSLDOIWNNTWMEWDREIDNTYTHLIYTBESQOQEKNOQELLDKXWASL 655

RESULT 5

US-10-190-305A-2

; Sequence 2, Application US/10190305A

; Publication No. US20030198621A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR

; FILE REFERENCE: 2302-18702 / 18702.002

; CURRENT APPLICATION NUMBER: US/10/190,305A

; CURRENT FILING DATE: 2002-07-05

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-305A-2

Query Match 83.0%; Score 2883.5; DB 12; Length 842;
Best Local Similarity 83.3%; Pred. No. 2e-252;
Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY 3 LWTVYVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 28 LWTVYVYGVVWKEATTTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 87
QY 63 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTETLSIIIVWQ 122
Db 88 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 141
QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVERIDDKNNTNTNTKYBLINCNTSVITQ 182

```
Db 142 MDRGEIKNCSFKVTTIRNMQKEALFYKLDVWPIDN-----DNTSVKLINCNSVITQ 196
Qy 183 ACPEVSEPIPIHYCTPTGFAILLKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
Db 197 ACPEVSEPIPIHYCAPAGFAILLKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 256
Qy 243 SLABEEVVRSENFTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYITGEILG 302
Db 257 SLABEGVVRSENFTNNAKTIIVQLKESVEINCTRPNNHTRKSTITIGRAFYATGDIIG 316
Qy 303 NIROAHCNISRAQWNNLTQIATTLREOFNGKTIARFQSSGGDPEIVMHSFNCGGERFYC 362
Db 317 DIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKQSSGGDPEIVMHSFNCGGERFYC 376
Qy 363 NSTQLFNSAMNVT-----SNGTWSVTRKQKOTGDIITLPCRKIQIINRWQVVGKAMYALP 417
Db 377 NSTQLFNSAMNVTGPNNTNGT-----ITLPCRKIQIINRWQVVGKAMYAPP 423
Qy 418 IKGLIRCSSNITGILLTRDGGGE-NQTEIPRPGGDMRDNRSELKYKVKVKEPLGVA 476
Db 424 IRGQIRCSSNITGILLTRDGGGEISNTTEIPRPGGDMRDNRSELKYKVKVKEPLGVA 483
Qy 477 PTKAKRRVVOREKRAVGMGLGAMFLGFLGAAGSTWGATSMALTVOARQLLSGIVQOQNLL 536
Db 484 PTKAKRRVVOREKRAV-TLGAFLGFLGAAGSTWGASLTTLTVQARQLLSGIVQOQNLL 542
Qy 537 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFGWCSGKLICTTAVPWNASWSN 596
Db 543 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFGWCSGKLICTTAVPWNASWSN 602
Qy 597 KTLQIWNMTWMEWDREIDNYTHLIYTLIBESQKQKQOELLQDKWASL 649
Db 603 KSLDQIWNMTWMEWDREIDNYTHLIYTLIBESQKQKQOELLQDKWASL 655

RESULT 6
US-09-476-242-2
; Sequence 2, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-2

Query Match 83.0%; Score 2883.5; DB 10; Length 847;
Best Local Similarity 83.3%; Pred. No. 2e-252;
Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

Qy 3 LWVTYYGVVPWKKEATTLFCASDAKAYDTEAHNWNATHACVPTNPQEVVLENVTENF 62
Db 33 LWVTYYGVVPWKKEATTLFCASDAKAYDTEAHNWNATHACVPTNPQEVVLENVTENF 92
Qy 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELISIVVWEQ 122
Db 93 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN-----WKE 146
Qy 123 RKGEMRNCSEFNITTSIRDKQREYALFYKLDVPEIDNKNNTNNTKYRLINCNTSVITQ 182
Db 147 MDRGEIKNCSFKVTTIRNMQKEALFYKLDVWPIDN-----DNTSVKLINCNSVITQ 201
Qy 183 ACPEVSEPIPIHYCTPTGFAILLKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
```

```
Db 202 ACPEVSEPIPIHYCAPAGFAILLKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 261
Qy 243 SLABEEVVRSENFTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYITGEILG 302
Db 262 SLABEGVVRSENFTNNAKTIIVQLKESVEINCTRPNNHTRKSTITIGRAFYATGDIIG 321
Qy 303 NIROAHCNISRAQWNNLTQIATTLREOFNGKTIARFQSSGGDPEIVMHSFNCGGERFYC 362
Db 322 DIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKQSSGGDPEIVMHSFNCGGERFYC 381
Qy 363 NSTQLFNSAMNVT-----SNGTWSVTRKQKOTGDIITLPCRKIQIINRWQVVGKAMYALP 417
Db 382 NSTQLFNSAMNVTGPNNTNGT-----ITLPCRKIQIINRWQVVGKAMYAPP 428
Qy 418 IKGLIRCSSNITGILLTRDGGGE-NQTEIPRPGGDMRDNRSELKYKVKVKEPLGVA 476
Db 429 IRGQIRCSSNITGILLTRDGGGEISNTTEIPRPGGDMRDNRSELKYKVKVKEPLGVA 488
Qy 477 PTKAKRRVVOREKRAVGMGLGAMFLGFLGAAGSTWGATSMALTVOARQLLSGIVQOQNLL 536
Db 489 PTKAKRRVVOREKRAV-TLGAFLGFLGAAGSTWGASLTTLTVQARQLLSGIVQOQNLL 547
Qy 537 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFGWCSGKLICTTAVPWNASWSN 596
Db 548 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFGWCSGKLICTTAVPWNASWSN 607
Qy 597 KTLQIWNMTWMEWDREIDNYTHLIYTLIBESQKQKQOELLQDKWASL 649
Db 608 KSLDQIWNMTWMEWDREIDNYTHLIYTLIBESQKQKQOELLQDKWASL 660

RESULT 7
US-10-032-162-13
; Sequence 13, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUEKE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-13

Query Match 81.4%; Score 2827.5; DB 15; Length 643;
Best Local Similarity 82.2%; Pred. No. 1.6e-247;
Matches 532; Conservative 40; Mismatches 56; Indels 19; Gaps 4;

Qy 3 LWVTYYGVVPWKKEATTLFCASDAKAYDTEAHNWNATHACVPTNPQEVVLENVTENF 62
Db 4 LWVTYYGVVPWKKEATTLFCASDAKAYDTEAHNWNATHACVPTNPQEVVLENVTENF 63
Qy 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELISIVVWEQ 122
Db 64 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTTNDSEGMT----- 117
Qy 123 RKGEMRNCSEFNITTSIRDKQREYALFYKLDVPEIDNKNNTNNTKYRLINCNTSVITQ 182
Db 118 -ERGIKNCSEFNITTSIRDEVOKEALFYKLDVWPIDN-----NNTSVKLISCDSVITQ 171
Qy 183 ACPEVSEPIPIHYCTPTGFAILLKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
```

```
Db 172 ACPISEPIPIHYCAPAGFAILACNDKTFNGKPCVKNSVQCTHGRPVVSTQLLNG 231
Qy 243 SLAEEVWIRSENFTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGELG 302
Db 232 SLAEEVWIRSDNFTNAKTIIVQLKESVEINCTRPNNHTRKSIHIGPRAFYTTGELG 291
Qy 303 NIROAHNCISRAQNNLTLOQIATTLREQFGNKTIAFNOSGGDPEIVMHSNCGGEFFYC 362
Db 292 DIRAHNCISRAKNDTLKQIVIKLREQFNKTIVFNHSSGGDPEIVMHSNCGGEFFYC 351
Qy 363 NSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIOIINRWQVVGKAMAYALPIKGLI 422
Db 352 NSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIOIINRWQVVGKAMAYALPIKGLI 405
Qy 423 RCSSNITGLLTRDGGGNGQTTETIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 482
Db 406 RCSSNITGLLTRDGGGNGQTTETIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 465
Qy 483 RVVOREKRAVGMFLGFLGAAGSTMGTATSMALTVOARQLLSGIVQOQNNLLRA 542
Db 466 RVVOREKRAVGMFLGFLGAAGSTMGTATSMALTVOARQLLSGIVQOQNNLLRA 524
Qy 543 QHLLQTLTVMGKIQARILAVERYLKDQQLLGPWGCGSKLICCTTAVPWNASWSNKTLDQI 602
Db 525 QHLLQTLTVMGKIQARILAVERYLKDQQLLGPWGCGSKLICCTTAVPWNASWSNKTLDQI 584
Qy 603 WNNMTWMEWREIDNYTHLYTLIESQOQEKNOQELLQDKWASL 649
Db 585 WNNMTWMEWREIDNYTHLYTLIESQOQEKNOQELLQDKWASL 631

RESULT 8
US-09-476-242-1
; Sequence 1, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-1

Query Match 80.9%; Score 2809.5; DB 10; Length 856;
Best Local Similarity 82.6%; Pred. No. 1.1e-245;
Matches 538; Conservative 38; Mismatches 56; Indels 19; Gaps 9;
Qy 3 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNQVVLNVTENF 62
Db 34 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNQVVLNVTENF 93
Qy 63 NNNKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTELSIIVVWEQ 122
Db 94 NNNKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTSSGRIME- 150
Qy 123 RGKGMNRCNFNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTNNTKYRLINCVTSVITQ 182
Db 151 --KGEIKNCNFNISTSRGKVQKEYAFFYKLDIIPDN-----DTTSYKLTSCNTSVITQ 203
Qy 183 ACPKVSFEPIPIHYCTPTGFAILLKNDKFNKNGTCTNVTSTVQCTHGRIPVSTQLLNG 242
Db 204 ACPKVSFEPIPIHYCAPAGFAILACNDKTFNGKPCVKNSVQCTHGRIPVSTQLLNG 263
Qy 243 SLAEEVWIRSENFTNAKTIIVQLNVSVEINCTRPNNHTRKRVTL--GPRVWYTTGEL 300
```

```
Db 264 SLAEEVWIRSVNFTDNAKTIIVQLNVSVEINCTRPNNHTRKRIORGPGRFVTIGKI 323
Qy 301 LGNTRQAHNCISRAQNNLTLOQIATTLREQFG-NKTIAFNOSGGDPEIVMHSNCGGEF 359
Db 324 -GNMROAHNCISRAKNNLTQIASKLREQFGNKTIIFKOSSGGDPEIVTHSNCGGEF 382
Qy 360 FYCNSSTOLFNSAMNVTSGTWSVTRKQKDTG-DIITLPCRIOIINRWQVVGKAMAYALPI 418
Db 383 FYCNSSTOLFNSAMNVTSGTWSVTRKQKDTG-DIITLPCRIOIINRWQVVGKAMAYALPI 439
Qy 419 KGLRCSNITGLLTRDGGGNGQTTETIFRPGGDMRDNRSSELYKYKVKIEPLGVAPT 478
Db 440 SGQRCSNITGLLTRDGGGNGQTTETIFRPGGDMRDNRSSELYKYKVKIEPLGVAPT 499
Qy 479 KAKERVVQREKRAVGMFLGFLGAAGSTMGTATSMALTVOARQLLSGIVQOQNNLLRA 538
Db 500 KAKERVVQREKRAVGMFLGFLGAAGSTMGTATSMALTVOARQLLSGIVQOQNNLLRA 558
Qy 539 IKAOHLLQTLTVMGKIQARILAVERYLKDQQLLGPWGCGSKLICCTTAVPWNASWSNKT 598
Db 559 IEAQHLLQTLTVMGKIQARILAVERYLKDQQLLGPWGCGSKLICCTTAVPWNASWSNKS 618
Qy 599 LDQIWNMTWMEWREIDNYTHLYTLIESQOQEKNOQELLQDKWASL 649
Db 619 LEQIWNMTWMEWREIDNYTHLYTLIESQOQEKNOQELLQDKWASL 669

RESULT 9
US-10-196-515-11
; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

Query Match 80.5%; Score 2795.5; DB 15; Length 856;
Best Local Similarity 82.3%; Pred. No. 2e-244;
Matches 536; Conservative 38; Mismatches 58; Indels 19; Gaps 9;
Qy 3 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNQVVLNVTENF 62
Db 34 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNQVVLNVTENF 93
Qy 63 NNNKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTELSIIVVWEQ 122
Db 94 DMKNKNNVQEMHEDIISLWQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSGRIME- 150
Qy 123 RGKGMNRCNFNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTNNTKYRLINCVTSVITQ 182
Db 151 --KGEIKNCNFNISTSRGKVQKEYAFFYKLDIIPDN-----DTTSYKLTSCNTSVITQ 203
Qy 183 ACPKVSFEPIPIHYCTPTGFAILLKNDKFNKNGTCTNVTSTVQCTHGRIPVSTQLLNG 242
Db 204 ACPKVSFEPIPIHYCAPAGFAILACNDKTFNGKPCVKNSVQCTHGRIPVSTQLLNG 263
```


; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-3

Query Match 79.4%; Score 2757; DB 15; Length 726;
Best Local Similarity 81.3%; Pred. No. 4.8e-241;
Matches 529; Conservative 43; Mismatches 55; Indels 24; Gaps 9;
QY 3 LWTYYGVVWKEATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
DB 34 LWTYYGVVWKEATITLFCASDAKAYETEVHNVWATHACVPTNPQEVVLENTNF 93
QY 63 NWNKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVVWQ 122
DB 94 NWNKNDVQEMHEDIISLWQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150
QY 123 RKGEMRNCSEFNTTSIRKQVREYALFKLDVEPDDKNTNTNNKYRLNCTSVITQ 182
DB 151 --KGEIKNSFNISTSKRGKVKKEYAFFYKLDIIPDNDP----TSYLTSCNTSVITQ 203
QY 183 ACPKVSFEPIHYCTPTGTGALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
DB 204 ACPKVSFEPIHYCAPAGAILKCNKNTFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 263
QY 243 SLAEVWVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRTL--GPRVWYTTGEI 300
DB 264 SLAEVWVIRSENFNTNAKTIIVQLNTSVSEINCTKPNNTTRKIRIHRGPGRAFTVGKI 323
QY 301 LGNIROAHNCISRAQWNTLQIATTLREQPG-NKTIAFNOSGGDPEIVMHSFNGCGEF 359
DB 324 -GNMROAHNCISRAKWSNTLKQIASKLREQPGNKKTIIFKQSSGGDPEIVTHSFNGCGEF 382
QY 360 FYCNSQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVVGKAMVALPI 418
DB 383 FYCKSTQLFNSWTSGSNNT-----EGSDTITLPCRIKQIINRWVVGKAMVALPI 434
QY 419 KGLIRCSSNITGLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478
DB 435 SGOIRCSSNITGLLTRDGGNSNNESEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 494
QY 479 KAKRRVVOREKRAVGMFLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 495 KAKRRVVOREKRAVG-IGALFLGAGSTMGAAASALTVOARQLSGIVQOQNLLRA 553
QY 539 IKAQOHLQLTQVWGIKQARILAVERYLKQDQLLGFWSGKLICTTAVPWNASWNT 598
DB 554 IEAQOHLQLTQVWGIKQARILAVERYLKQDQLLGIWGSCKLICCTTAVPWNASWNT 613
QY 599 LDQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 649
DB 614 LEQIWNMTWMDREIDNTHYLSLHSLIEESQIQEENQELLELDKWASL 664

RESULT 12

US-10-196-515-12
; Sequence 12, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; LABRANCHE, Celia C.

; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-12

Query Match 79.3%; Score 2755; DB 15; Length 759;
Best Local Similarity 81.3%; Pred. No. 7.8e-241;
Matches 529; Conservative 43; Mismatches 55; Indels 24; Gaps 9;
QY 3 LWTYYGVVWKEATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
DB 34 LWTYYGVVWKEATITLFCASDAKAYETEVHNVWATHACVPTNPQEVVLENTNF 93
QY 63 NWNKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVVWQ 122
DB 94 NWNKNDVQEMHEDIISLWQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150
QY 123 RKGEMRNCSEFNTTSIRKQVREYALFKLDVEPDDKNTNTNNKYRLNCTSVITQ 182
DB 151 --KGEIKNSFNISTSKRGKVKKEYAFFYKLDIIPDNDP----TSYLTSCNTSVITQ 203
QY 183 ACPKVSFEPIHYCTPTGTGALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242
DB 204 ACPKVSFEPIHYCAPAGAILKCNKNTFNGTGPCTNVSTVQCTHGIRPVVSTQLLNG 263
QY 243 SLAEVWVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRTL--GPRVWYTTGEI 300
DB 264 SLAEVWVIRSENFNTNAKTIIVQLNTSVSEINCTKPNNTTRKIRIHRGPGRAFTVGKI 323
QY 301 LGNIROAHNCISRAQWNTLQIATTLREQPG-NKTIAFNOSGGDPEIVMHSFNGCGEF 359
DB 324 -GNMROAHNCISRAKWSNTLKQIASKLREQPGNKKTIIFKQSSGGDPEIVTHSFNGCGEF 382
QY 360 FYCNSQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVVGKAMVALPI 418
DB 383 FYCKSTQLFNSWTSGSNNT-----EGSDTITLPCRIKQIINRWVVGKAMVALPI 434
QY 419 KGLIRCSSNITGLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478
DB 435 SGOIRCSSNITGLLTRDGGNSNNESEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 494
QY 479 KAKRRVVOREKRAVGMFLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB 495 KAKRRVVOREKRAVG-IGALFLGAGSTMGAAASALTVOARQLSGIVQOQNLLRA 553
QY 539 IKAQOHLQLTQVWGIKQARILAVERYLKQDQLLGFWSGKLICTTAVPWNASWNT 598
DB 554 IEAQOHLQLTQVWGIKQARILAVERYLKQDQLLGIWGSCKLICCTTAVPWNASWNT 613
QY 599 LDQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 649
DB 614 LEQIWNMTWMDREIDNTHYLSLHSLIEESQIQEENQELLELDKWASL 664

RESULT 13

US-09-891-609-4
; Sequence 4, Application US/09891609
; Patent No. US2002012738A1
; GENERAL INFORMATION:

; APPLICANT: Stamatos, Leonidas
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,608
; FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; TYPE: PRF
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-4

Query Match 78.5%; Score 2728.5; DB 10; Length 619;
Best Local Similarity 79.6%; Pred. No. 1.5e-238;
Matches 520; Conservative 35; Mismatches 39; Indels 59; Gaps 8;
QY 3 LWTVVYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
Db 6 LWTVVYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 65
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSTIIIVWEQ 122
Db 66 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 119
QY 123 RKGEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYLINCNSTVITQ 182
Db 120 MDRGEIKNCSPKVGAG-----KLINCNSTVITQ 147
QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLLNG 242
Db 148 ACPKVSFPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNG 207
QY 243 SLAEEVVRISNFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302
Db 208 SLAEEVVRISNFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPRAFYATGDIIG 267
QY 303 NIROAHNCISRAQWNTLQIATTLREQFGNKTIAFNOSGGDPBIVMHSFNCGGEPFYC 362
Db 268 DIROAHNCISGEKNWNTLKQIVTKLQAFGNKTIIVFKOSSGGDPBIVMHSFNCGGEPFYC 327
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKOTGDIITLPCRKQIINRWQVVGKAMYALP 417
Db 328 NSTQLFNSAWNTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYAPP 374
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476
Db 375 IRGQIRCSNITGLLLTRDGGGEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 434
QY 477 PTKAKRRVQREKRAVGMFLGAGSTMGATSMALTVOARQLLSGI VQOQNNLL 536
Db 435 PTKAKRRVQREKRAV-TLGMFLGAGSTMGARSLTITVOARQLLSGI VQOQNNLL 493
QY 537 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWGCSGKLICTTAVPNWASWN 596
Db 494 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWGCSGKLICTTAVPNWASWN 552
QY 597 KTLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOELLQLDKWSL 649
Db 553 KSLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOELLQLDKWSL 605

RESULT 14
US-09-891-609-2
; Sequence 2, Application US/09891609
; Patent No. US20020127238A1
; GENERAL INFORMATION:
; APPLICANT: Stamatos, Leonidas
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR
; FILE REFERENCE: 2570-1-002N
; CURRENT APPLICATION NUMBER: US/09/891,609

; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,608
; FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 646
; TYPE: PRF
; ORGANISM: Human immunodeficiency virus type 1
US-09-891-609-2

Query Match 78.5%; Score 2728.5; DB 10; Length 646;
Best Local Similarity 79.6%; Pred. No. 1.6e-238;
Matches 520; Conservative 35; Mismatches 39; Indels 59; Gaps 8;
QY 3 LWTVVYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62
Db 33 LWTVVYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 92
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSTIIIVWEQ 122
Db 93 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 146
QY 123 RKGEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYLINCNSTVITQ 182
Db 147 MDRGEIKNCSPKVGAG-----KLINCNSTVITQ 174
QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLLNG 242
Db 175 ACPKVSFPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLLNG 234
QY 243 SLAEEVVRISNFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302
Db 235 SLAEEVVRISNFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPRAFYATGDIIG 294
QY 303 NIROAHNCISRAQWNTLQIATTLREQFGNKTIAFNOSGGDPBIVMHSFNCGGEPFYC 362
Db 295 DIROAHNCISGEKNWNTLKQIVTKLQAFGNKTIIVFKOSSGGDPBIVMHSFNCGGEPFYC 354
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKOTGDIITLPCRKQIINRWQVVGKAMYALP 417
Db 355 NSTQLFNSAWNTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYAPP 401
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476
Db 402 IRGQIRCSNITGLLLTRDGGGEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 461
QY 477 PTKAKRRVQREKRAVGMFLGAGSTMGATSMALTVOARQLLSGI VQOQNNLL 536
Db 462 PTKAKRRVQREKRAV-TLGMFLGAGSTMGARSLTITVOARQLLSGI VQOQNNLL 520
QY 537 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWGCSGKLICTTAVPNWASWN 596
Db 521 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWGCSGKLICTTAVPNWASWN 579
QY 597 KTLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOELLQLDKWSL 649
Db 580 KSLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOELLQLDKWSL 632

RESULT 15
US-10-032-162-17
; Sequence 17, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUELKE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADSON J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162

Job time : 39 secs

```
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
;
US-10-032-162-17
```

```
Query Match      78.3%; Score 2720.5; DB 15; Length 625;
Best Local Similarity 80.1%; Pred. No. 7.9e-238;
Matches 518; Conservative 36; Mismatches 56; Indels 37; Gaps 5;

.
QY      3 LWTYYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVWLENVTENF 62
Db      4 LWTYYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVWLENVTENF 63
QY      63 NMKNMNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTLNNTNTTNTTIELSIIVVWEQ 122
Db      64 NMKNMNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGM----- 117
QY      123 RGKGMNRCNFNTTSTRDVQREYALFYKLDVEPIDDNKNTNTNTKYRLINCNTSVITQ 182
Db      118 -ERGEIKNCSFNITTSRDEQVEYALFYKLDVXXIDN-----NNTSYRLISCOTSVITQ 171
QY      183 ACPKVSPEPIPIHYCTPTGALLKCNCKKFGTGTNNVSTVQCTHGIKIRPVVSTQLLLNG 242
Db      172 ACPKISPEPIPIHYCAPAGFAILLKNDKTFNGKXPKCNVSTXCTHGIKIRPVVSTQLLLNG 231
QY      243 SLAESEVVISENFTNNAKTIIVQLNVSVIEINCTRPNNHTRKVTLGPRVWYTTGEILG 302
Db      232 SLAESEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNNG-----AG 273
QY      303 NIRQAHCNISRAQWNTLQOIATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGRFFYC 362
Db      274 DIRQAHCNISRAKNDTLKQIVIKLREQFNKTIIVNHSGGDPEIVMHSFNCGRFFYC 333
QY      363 NSTQLNSANVTSGNFWSTRKQKDTGDIITLPCRKQIINRWQVVGVKAMYALPIKGLI 422
Db      334 NSTQLFNSWNNTSGS-----NNTGNTITLPCRKQIINRWQEVGVKAMYAPPINGQI 387
QY      423 RCSSNITGLLLTRDGGNGOTTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKR 482
Db      388 RCSSNITGLLLTRDGGNGENGTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTCKR 447
QY      483 RVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQQONNLLRAIKQ 542
Db      448 RVVQREKRAVG-IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQONNLLRAIEAQ 506
QY      543 QHLLQLTVWGIGKQIARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWSNKTLDQI 602
Db      507 QRMQLTVWGIGKQIARVLAVERYLGDQQLLGIWCGSGKLICTTAVPWNASWSNKSIDRI 566
QY      603 WNNMTWWEWDREIDNTHLYTLIERSONQOEKXQQLQDLQKWSL 649
Db      567 WNNMTWWEWEREIDNVTSELYTLIERSONQOEKXQQLQDLQKWSL 613
```

Search completed: January 13, 2004, 09:16:34

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:06:27 ; Search time 40 Seconds

(without alignments)
4186.903 Million cell updates/sec

Title: US-09-938-406-1_COPY_33_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVFWKREATT.....QNQQEKXQQLLDKWL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2908	83.7	854	15	O40222 human immun
2	2896.5	83.4	853	15	O40222 human immun
3	2889	83.2	841	15	O9QKJ1 human immun
4	2872.5	82.7	863	15	O77989 human immun
5	2869.5	82.6	843	15	O70150 human immun
6	2865.5	82.5	851	15	O56110 human immun
7	2865.5	82.5	863	15	O42031 human immun
8	2862.5	82.4	863	15	O9WJU8 human immun
9	2860	82.3	854	15	O9WJU8 human immun
10	2858.5	82.3	868	15	O9WJU5 human immun
11	2857	82.2	858	15	O9WJU4 human immun
12	2855.5	82.2	853	15	O9WJU2 human immun
13	2855.5	82.2	855	15	O9WJU2 human immun
14	2854	82.2	858	15	O9WJU2 human immun
15	2853	82.1	864	15	O9WJU2 human immun
16	2853	82.1	864	15	O9WJU2 human immun

17	2852	82.1	864	15	O9YP39 human immun
18	2850	82.0	856	15	O9EIR8 human immun
19	2849.5	82.0	847	15	O75760 human immun
20	2846.5	81.9	871	15	O902L4 human immun
21	2845.5	81.9	860	15	O9EIS7 human immun
22	2845	81.9	856	15	O77694 human immun
23	2844.5	81.9	853	15	O91023 human immun
24	2844.5	81.9	855	15	O8Q367 human immun
25	2844	81.9	856	15	O9EIS5 human immun
26	2844	81.9	856	15	O72993 human immun
27	2843.5	81.9	857	15	O89654 human immun
28	2843.5	81.9	864	15	O9YP31 human immun
29	2842	81.8	854	15	O78225 human immun
30	2841.5	81.8	853	15	O56108 human immun
31	2841	81.8	858	15	O9DL80 human immun
32	2840.5	81.8	849	15	O8Q851 human immun
33	2840.5	81.8	853	15	O9EIS9 human immun
34	2840.5	81.8	859	15	P87924 human immun
35	2840	81.8	858	15	O9DL79 human immun
36	2838	81.7	848	15	O9EIS3 human immun
37	2838	81.7	854	15	O99C08 human immun
38	2838	81.7	860	15	O9YP50 human immun
39	2838	81.7	862	15	O9EIS2 human immun
40	2837	81.7	852	15	O92761 human immun
41	2837	81.7	852	15	O41883 human immun
42	2836.5	81.6	852	15	O73303 human immun
43	2835.5	81.6	851	15	O8Q852 human immun
44	2835.5	81.6	861	15	O9WJV5 human immun
45	2835	81.6	735	15	O9IW16 human immun

ALIGNMENTS

RESULT 1

O40222 PRELIMINARY; PRT; 854 AA.

AC O40222; (TREMREL. 05, Created)
 DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
 DT 01-OCT-2002 (TREMREL. 22, Last annotation update)
 DE Env polyprotein.
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RC STRAIN=AD8;
 RC STRAIN=AD8;
 RX MEDLINE=96432129; PubMed=8835195;
 RA Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
 RA Martin M.A., Peden K.W.;
 RT "Construction and characterization of a stable full-length macrophage-tropic HIV type 1 molecular clone that directs the production of high titers of progeny virions.";
 RT Aids Res. Hum. Retroviruses 12:191-194(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=AD8;
 RC Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,
 RA Martin M.A., Peden K.W.C.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF004394; AAB64170.1;
 DR InterPro; IPR00328; Env GP41.
 DR InterPro; IPR00777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 854 AA; 97291 MW; 06C45E9103C6C12 CRC64;

Query Match 83.7%; Score 2908; DB 15; Length 854;
 Best Local Similarity 84.0%; Pred. No. 1.3e-230;
 Matches 548; Conservative 36; Mismatches 48; Indels 20; Gaps 7;


```

DR EMBL: AF112547; AAF13325.1; -.
DR InterPro: IPR000328; Env GP41.
DR Pfam: PF00516; GP120.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 841 AA; 95389 MW; E9086659E7E1D447 CRC64;

Query Match      83.2%; Score 2889; DB 15; Length 841;
Best Local Similarity 83.5%; Pred. No. 4.6e-229;
Matches 542; Conservative 38; Mismatches 53; Indels 16; Gaps 4;

QY 3 LWTVTYGVVPWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPPOEVVLENTENF 62
DB 20 LWTVTYGVVPWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPPOEVVLENTENF 79
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWEO 122
DB 80 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWEO 135
QY 123 RKGEMRKCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 182
DB 136 EMKGMKNCSEFNITTSIRDKVQREYALFYKLDVPIGSD---SDNTSYRLINCNTSVITQ 192
QY 183 ACPKVSFEPIPHYCTPTGPFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
DB 193 ACPKVSFEPIPHYCTPTGPFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 252
QY 243 SLAEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWVTGTEIIG 302
DB 253 SLAEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWVTGTEIIG 312
QY 303 NTRQACNISRQWNTTLOQIATTLREQFGNKTIAFNQSGGDPPEIVMHSFNGGGEFFYC 362
DB 313 DIRQACNLSRTQWNTLRFQIKLREQFGNKTIAFNQSGGDPPEIVMHSFNGGGEFFYC 372
QY 363 NSTQIFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMVALPIKGLI 422
DB 373 NSTQIFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMVALPIKGLI 425
QY 423 RCSSNITGLLTRDGGENQOT--TEIFRPGGDMRDNRWSELYKYKVKIIEPLGVAPTKA 480
DB 426 RCTSNITGLLTRDGGNNNSNETETFRPGGDMRDNRWSELYKYKVKIIEPLGVAPTKA 485
QY 481 KRVVQREKRAVGMGLGAMFLGAGSTMGATSMALTVAQRLSGIVQQNNLRAIK 540
DB 486 KRVVQREKRAVGMGLGAMFLGAGSTMGATSMALTVAQRLSGIVQQNNLRAIE 545
QY 541 AQOHLQLTVWGIKQIARILAVERYLKDOQLLGFWCGSGKLICTTAVPWNASWNSKTL 600
DB 546 AQOHLQLTVWGIKQIARILAVERYLKDOQLLGFWCGSGKLICTTAVPWNASWNSK 605
QY 601 QIWNNTMWEDEIDNYTHLYTLIEESQKQKQELLQDLKWSL 649
DB 606 KIWNNTMWEDEIDNYTHLYTLIEESQKQKQELLQDLKWSL 654

RESULT 4
Q77989
ID Q77989 PRELIMINARY; PRT; 863 AA.
AC Q77989;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF128A;

```

```

RX MEDLINE=91056585; PubMed=2243391;
RA Liu Z.-Q.; Wood C.; Levy J.A.; Cheng-Mayer C.;
RT "The viral envelope gene is involved in macrophage tropism of a human
immunodeficiency virus type 1 strain isolated from brain tissue.";
RL J. Virol. 64:6148-6153(1990).
DR EMBL: M95292; AAA44331.1; -.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 863 AA; 98222 MW; 0157397FA8B4C5D7 CRC64;

Query Match      82.7%; Score 2872.5; DB 15; Length 863;
Best Local Similarity 82.6%; Pred. No. 1.1e-227;
Matches 543; Conservative 43; Mismatches 48; Indels 23; Gaps 7;

QY 3 LWTVTYGVVPWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPPOEVVLENTENF 62
DB 33 LWTVTYGVVPWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPPOEVVLENTENF 92
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCT---DLNNTNTTNTLS 115
DB 93 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNDTNTND--TRATISS 151
QY 116 IIVWEOQK--GEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNNTKYRLIN 174
DB 152 L-----RGTEIKNCSEFNITTSIRDKVQREYALFYKLDIVFINN-----NNITYRLIN 200
QY 175 CNTSVITQACPKVSFEPIPHYCTPTGPFALLKCNKKFNGTGPCTNVSTVQCTHGIRPV 234
DB 201 CNTSVITQACPKVSFEPIPHYCTPTGPFALLKCNKKFNGTGPCTNVSTVQCTHGIRPV 260
QY 235 STQLLNGSLAEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVW 294
DB 261 STQLLNGSLAEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVW 320
QY 295 YTGILNIRQAHNCNISRQWNTTLOQIATTLREQFGNKTIAFNQSGGDPPEIVMHSFN 354
DB 321 YTGILNIRQAHNCNISRQWNTTLOQIATTLREQFGNKTIAFNQSGGDPPEIVMHSFN 380
QY 355 CGGEFFYCNSTOLFNSAWNTS--NGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKA 412
DB 381 CGGEFFYCNSTOLFNSAWNTS--NGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKA 440
QY 413 MYALPIKGLIRCSNITGLLTRDGGENQOTTEIFRPGGDMRDNRWSELYKYKVKIIEP 472
DB 441 MYALPIKGLIRCSNITGLLTRDGGENQOTTEIFRPGGDMRDNRWSELYKYKVKIIEP 500
QY 473 LGVAPTKARVVRQREKRAVGMGLGAMFLGAGSTMGATSMALTVAQRLSGIVQQ 532
DB 501 LGVAPTKARVVRQREKRAVGMGLGAMFLGAGSTMGATSMALTVAQRLSGIVQQ 559
QY 533 NNLLRAIKAQOHLQLTVWGIKQIARILAVERYLKDOQLLGFWCGSGKLICTTAVPWN 592
DB 560 NNLLRAIKAQOHLQLTVWGIKQIARILAVERYLKDOQLLGFWCGSGKLICTTAVPWN 619
QY 593 SWSNKTLDQIWNNTMWEDEIDNYTHLYTLIEESQKQKQELLQDLKWSL 649
DB 620 SWSNKTLDQIWNNTMWEDEIDNYTHLYTLIEESQKQKQELLQDLKWSL 676

RESULT 5
Q70150
ID Q70150 PRELIMINARY; PRT; 843 AA.
AC Q70150;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.

```



```
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;
RA "Genetic variation of HIV type 1 in four World Health Organization-
RT sponsored vaccine evaluation sites: generation of functional envelope
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,
RT and E. WHO Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1359-1368 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194691; PubMed=7545977;
RA WHO Global Programme on AIDS;
RA "HIV type 1 variation in World Health Organization-sponsored vaccine
RT evaluation sites: genetic screening, sequence analysis, and
RT preliminary biological characterization of selected viral strains. WHO
RT Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1327-1343 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RA "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX Allen E.E.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U088801; AAC05185.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON TER
SQ SEQUENCE 843 AA; 95679 MW; 14DB49A6D2F5FCE4 CRC64;

Query Match 82.6%; Score 2869.5; DB 15; Length 843;
Best Local Similarity 83.4%; Pred. No. 1.8e-227;
Matches 542; Conservative 39; Mismatches 42; Indels 27; Gaps 8;

QY 3 LWVTYYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 62
DB 31 LWVTYYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 90
QY 63 NWNKNNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTNTTSLIIVWE 121
DB 91 NWNKNNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTSSTME----- 142
QY 122 QRGKEMNCNFIITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVIT 181
DB 143 ---GGEIKNCNFIITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVIT 193
QY 182 QACPKVSEPIPIHYCTPTGFPALLKCNCKKNGTGPCNTNSTVQCTHGIRPVWSTQLLN 241
DB 194 QACPKVSEPIPIHYCTPTGFPALLKCNCKKNGTGPCNTNSTVQCTHGIRPVWSTQLLN 253
QY 242 GSLAEVEEVVISENTNNAKTIIVQLNVSVEINCTRNPNHTRKRVTLGPGRVWVTTGIL 301
DB 254 GSLAEVEEVVISENTNNAKTIIVQLNVSVEINCTRNPNHTRKRVTLGPGRVWVTTGIL 313
QY 302 GNIRQAHCNISRAQWNNTLQIATTLREQFGNKTIAFNCSSGGDPEIWMHSFNCGGEFFY 361
```

```
DB 314 GDIRQAHCNLSSTKWNNTLRQITEKLRQFGNKTIIVNCSSGGDPEIWMHSFNCGGEFFY 373
QY 362 CNSTQLFNSAWNTVTSNGTWSVTRKQKDTGD-IITLPCRIKQIINRWQVVGKAMVAPLTKG 420
DB 374 CNTQLFNSWNTDTS--TWN-----NNTNGTITLPCRIKQIINRWQVVGKAMVAPLTKG 426
QY 421 LIRCSSNITGLLLTRDGG-GENQTEIFRPGGGMDNRNRSLEYKYKVKVTEPLGVAPTK 479
DB 427 QIRCSSNITGLLLTRDGGSENKITEIFRPGGGMDNRNRSLEYKYKVKVTEPLGVAPTK 486
QY 480 AKREVQREKRAVGMGLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAI 539
DB 487 PKREVQREKRAVGTIGAMFLGFLGAGSTMGASITLTVQARQLLSGIVQQNNLLRAI 546
QY 540 KAOQHLLQLTWGIGKQIQAARILAVERYLKQQLLGFWGCSGKLICTTAVPWNASWNTKL 599
DB 547 EAOQHLLQLTWGIGKQIQAARILAVERYLKQQLLGFWGCSGKLICTTAVPWNASWNTKL 606
QY 600 DOIWNNTWMDREIDNYTHLYTLIEESQOQEKQOELLQDKWASL 649
DB 607 DKIWNTWMDREIDNYTHLYTLIEESQOQEKQOELLQDKWASL 656

RESULT 6
O56110 PRELIMINARY; PRT; 851 AA.
AC O56110;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SFVHS8.1;
RX MEDLINE=98178716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RA "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337 (1998).
DR EMBL; AF025756; AAC40593.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDAC83EB6464531 CRC64;

Query Match 82.5%; Score 2865.5; DB 15; Length 851;
Best Local Similarity 83.2%; Pred. No. 4e-227;
Matches 538; Conservative 42; Mismatches 52; Indels 15; Gaps 5;

QY 3 LWVTYYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 62
DB 33 LWVTYYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 92
QY 63 NWNKNNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTSLIIVWEQ 122
DB 93 NWNKNNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTSSS-----GET 146
QY 123 RGKEMNCNFIITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVITQ 182
DB 147 MERGEIKNCNFIITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVITQ 200
QY 183 ACPKVSFEPIPIHYCTPTGFPALLKCNCKKNGTGPCNTNSTVQCTHGIRPVWSTQLLN 242
```

```
Db 201 ACPKVSFPIPIHYCAPAGFAILKCNDRKPNKGCTNVTSTVQCTHGIRPVVSTQPLNG 260
Qy 243 SLAEVWIRSENFNTNNAKIIIVQLNVSVSEINCTRPNNHTRKRVLTGPRGVWYTTGEILG 302
Db 261 SLAEVWIRSENFNTNNAKIIIVQLNVSVSEINCTRPNNHTRKRVLTGPRGVWYTTGEILG 320
Qy 303 NTRQAHNISRAQWNTLQOIIATTLREQFG-NKTIAFNQSSGGDPPIVHSHFNCGGEFFYC 362
Db 321 DIRQAHNLSEAKWHTLEQIAKRLREQFGNKTIIIVFNQSSGGDPPIVHSHFNCGGEFFYC 380
Qy 363 NSTQFNLSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 422
Db 381 NSTKLFNRTSV--NSTWMDTEGVNTGENITLPCRIKQIINRWQVGVKAMVALPIKGLI 438
Qy 423 RCSSNITGLLTRDGGNGNQTETPRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKR 482
Db 439 RCSSNITGLLTRDGGNSNETETPRPGGDMRDNRSELYKYKVKVIEPLGVAPTAKR 498
Qy 483 RVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRAIKAO 542
Db 499 RVVQREKRAVG-IGAVFLGFLGAAGSTMGASMTLTVOARQLLSGIVQOQNLLRAIEAQ 557
Qy 543 QHLLQLTVMGIIKQOARILAVERYLKDQQLIGFWGCSGKLICTTAVPWNASNSKTLDOI 602
Db 558 QHLLQLTVMGIIKQOARILAVERYLKDQQLIGFWGCSGKLICTTAVPWNASNSKTLDOI 617
Qy 603 WNNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 649
Db 618 WNNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 664

RESULT 7
O42031
ID O42031 PRELIMINARY; PRT; 863 AA.
AC O42031;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE Envelope polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ENL4-3;
RA Fang G., Weiser B., Visosky A., Burger H.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003888; AAB64286.1; -.
DR EMBL; AF003887; AAB64278.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 863 AA; 98166 MW; 918DD40F4F0C16E7 CRC64;

Query Match 82.5%; Score 2865.5; DB 15; Length 863;
Best Local Similarity 82.3%; Pred. No. 4.1e-227;
Matches 540; Conservative 40; Mismatches 53; Indels 23; Gaps 5;

Qy 4 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENFN 63
Db 34 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENFN 93
Qy 64 MWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD-----LNTNNTTNTTSLI 116
Db 94 MWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDNLKNATVNNANNTNNS- 149
Qy 117 IVVWEORGKGMENCSFNITSTRDKVQREYALFYKLDVEPID--DNKNTTNTKYLRLN 174
Db 150 ---WEKWEKEIKNCSEFNITSTRDKVQREYALFYKLDVVPIDNANNATNTYTSLIS 206
```

```
Qy 175 CNTSVITQACPKVSFPIPIHYCTPTGFALLKCNDRKPNKGCTNVTSTVQCTHGIRPVV 234
Db 207 CNTSVITQACPKVSFPIPIHYCAPAGFAILKCNDRKPNKGCTNVTSTVQCTHGIRPVV 266
Qy 235 STOLLNGLSLAEVWIRSENFNTNNAKIIIVQLNVSVSEINCTRPNNHTRKRVLTGPRGVW 294
Db 267 STOLLNGLSLAEVWIRSENFNTNNAKIIIVQLNVSVSEINCTRPNNHTRKRVLTGPRGVW 326
Qy 295 YTTGELGNIRQAHCNISRAQWNTLQOIIATTLREQFG-NKTIAFNQSSGGDPPIVHSHF 353
Db 327 YTTGDIIGDIRQAHCNVSRAKNNTLVKIVKLEQFGHNKTIVFNHSSGGDLIVTHSF 386
Qy 354 NCGGFFVYCNSTOLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAM 413
Db 387 ICGGFFVYCNSTOLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAM 440
Qy 414 YALPIKGLIRCSSNITGLLTRDGGNGNQTETPRPGGDMRDNRSELYKYKVKVIEPL 473
Db 441 YAPPTRGQIRCLSNITGLLTRDGGDTNGTEVFRPGGDMRDNRSELYKYKVKVIEPL 500
Qy 474 GVAPTAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQN 533
Db 501 GVAPTAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGASMTLTVOARQLLSGIVQOQN 560
Qy 534 NLLRAIKAQHLLQLTVMGIIKQOARILAVERYLKDQQLIGFWGCSGKLICTTAVPWNAS 593
Db 561 NLLRAIEAQHLLQLTVMGIIKQOARILAVERYLKDQQLIGFWGCSGKLICTTAVPWNAS 620
Qy 594 WSKNTLDOIWNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 649
Db 621 WSKNSLEIWNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 676

RESULT 8
O9WJ08
ID O9WJ08 PRELIMINARY; PRT; 863 AA.
AC O9WJ08;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE Envelope polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,
RA Burger H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69590; AAD10923.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 863 AA; 98148 MW; 5275ACD8476BAE02 CRC64;

Query Match 82.4%; Score 2862.5; DB 15; Length 863;
Best Local Similarity 82.3%; Pred. No. 7.2e-227;
Matches 540; Conservative 39; Mismatches 54; Indels 23; Gaps 5;

Qy 4 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENFN 63
Db 34 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENFN 93
Qy 64 MWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD-----LNTNNTTNTTSLI 116
Db 94 MWKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDNLKNATVNNANNTNNS- 149
```

```

QY 117 IUVWEQKRGEMRNCSENIITTSIRDKVQREYALFYKLDVEPID--DNKNTTNTKYRLIN 174
Db 150 ---WEKEKGKNCSENIITTSIRDKVQREYALFYKLDVEPIDNANNNAINTSYRLIS 206
QY 175 CNTSVITQACPKVSFEPIPIHYCTPTGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 234
Db 207 CNTSVITQACPKVSFEPIPIHYCAPAGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 266
QY 235 STQLLNGSLAEBEVEVIRSENFNTNNAKTIIVQLNVSVIEINCTRPNNHTRKRVITLPGRVW 294
Db 267 STQLLNGSLAEBEVEVIRSENFNTNNAKTIIVQLNVSVIEINCTRPNNHTRKRVITLPGRVW 326
QY 295 YTTGEILGNIRQAHNCISRAQWNNTLQOIATTLRBOFG-NKTIAPNOSGGDPEIVMHSF 353
Db 327 YTTGDIIGIRQAHNCISRAQWNNTLQOIVKLEKBOFGHNKTIIVFNHSSGGDLEIVHSP 386
QY 354 NCGGEFFYCSNTQSFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAM 413
Db 387 ICGGEFFYCSNTQSFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAM 440
QY 414 YALPIKGLIRCSSNITGLLLTRDGGENQTTETPRPGGDMRDNRSELKYKYVVKIEPL 473
Db 441 YAPPIRGQIRCLSNITGLLLTRDGGDTENGTEVFRPGGDMRDNRSELKYKYVVKIEPL 500
QY 474 GVAPTKAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 533
Db 501 GVAPTKAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 560
QY 534 NLLRAIKAAQHLQLTVMGKIQARILAVERYLKDOQLLFGWCGSGKLICTTAVPNWAS 593
Db 561 NLLRAIKAAQHLQLTVMGKIQARILAVERYLKDOQLLFGWCGSGKLICTTAVPNWAS 620
QY 594 WSKNTLQIWNNTWMEWDREIDNYTHLYTLIBESQKQKQELLOLDKWSL 649
Db 621 WSKSLNEIWNNTWMEWDREIDNYTHLYTLIBESQKQKQELLOLDKWSL 676

RESULT 9
QY 09E1S8 PRELIMINARY; PRT; 854 AA.
AC Q9E1S8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Jack N., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDaniel C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RT Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF277057; ANG22502.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97210 MW; 2DBA931678837491 CRC64;

Query Match 82.3%; Score 2860; DB 15; Length 854;
Best Local Similarity 81.9%; Pred. No. 1.1e-226;
Matches 535; Conservative 52; Mismatches 46; Indels 20; Gaps 7;

QY 1 ANLWTVYGVVPMKEATTTTFCASDAKAYDTEAHNVWATHACVPTDNPQEVVLENVTE 60
Db 31 ANLWTVYGVVPMKEATTTTFCASDAKAYDTEAHNVWATHACVPTDNPQEVVLENVTE 90

```

```

QY 61 NFNMKNMVEQMEHDIISLWDSIKPCVKLTPLCVTLNCTDLNTNT-TNTTSLIIV 119
Db 91 NFNMKNMVEQMEHDIISLWDSIKPCVKLTPLCVTLNCTDLNTNT-TNTTSLIIV 144
QY 120 WEQKKGEMRNCSENIITTSIRDKVQREYALFYKLDVEPIDDN--KNTTNTKYRLIN 177
Db 145 ---GKEIKNCSENIITTAVERDKYKTYALFYRLDVPIDDDHGNSSSNKYRLIN 200
QY 178 SVITQACPKVSFEPIPIHYCTPTGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 237
Db 201 SVITQACPKVSFEPIPIHYCAPAGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 260
QY 238 LLNGSLAEBEVEVIRSENFNTNNAKTIIVQLNVSVIEINCTRPNNHTRKRVITLPGRV 297
Db 261 LLNGSLAEBEVEVIRSENFNTNNAKTIIVQLNVSVIEINCTRPNNHTRKRVITLPGRV 319
QY 298 GEILGNIRQAHNCISRAQWNNTLQOIATTLRBOFGNKTIAFNQSSGGDPEIVMHSF 357
Db 320 GEILGNIRQAHNCISRAQWNNTLQOIVKLEKBOFGHNKTIIVFNHSSGGDPEIVMHSF 379
QY 358 EFFYCSNTQSFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAM 417
Db 380 EFFYCSNTQSFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAM 434
QY 418 IKGLIRCSSNITGLLLTRDGGENQTTETPRPGGDMRDNRSELKYKYVVKIEPL 476
Db 435 IRGQIRCSSNITGLLLTRDGGTNNGTEVFRPGGDMRDNRSELKYKYVVKIEPL 494
QY 477 PTKAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 536
Db 495 PTKAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 554
QY 537 RAIKAAQHLQLTVMGKIQARILAVERYLKDOQLLFGWCGSGKLICTTAVPNWAS 596
Db 555 RAIKAAQHLQLTVMGKIQARILAVERYLKDOQLLFGWCGSGKLICTTAVPNWAS 614
QY 597 KTLQIWNNTWMEWDREIDNYTHLYTLIBESQKQKQELLOLDKWSL 649
Db 615 KSLSDIWNNTWMEWDREIDNYTHLYTLIBESQKQKQELLOLDKWSL 667

RESULT 10
QY 09WJW5 PRELIMINARY; PRT; 868 AA.
AC Q9WJW5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,
RA Burger H.;
RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to
RT Non-Progressive Infection.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69584; AAD10875.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 868 AA; 98569 MW; F2D138F0F8BFFC02 CRC64;

Query Match 82.3%; Score 2860; DB 15; Length 868;
Best Local Similarity 82.6%; Pred. No. 1.2e-226;
Matches 543; Conservative 36; Mismatches 60; Indels 18; Gaps 7;

```

Best Local Similarity 81.7%; Pred. No. 1.5e-226;
Matches 537; Conservative 44; Mismatches 53; Indels 23; Gaps 5;

QY 3 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62
DB 33 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92
QY 63 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 118
DB 93 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 149
QY 119 VMEQKGMKRNCSFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNTKYLINC 175
DB 150 -WGMERGEIKNGSFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNTKYLINC 208
QY 176 NTSVITQACPKVSFEPIPIHYCTPTGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPVV 325
DB 209 NTSVITQACPKVSFEPIPIHYCAPAGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPVV 368
QY 236 TQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTNRKRVTLGPGRVMY 295
DB 269 TQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTNRKRVTLGPGRVMY 328
QY 296 TTGEILGNTRQAHNCNISRAQWNTLQOIATTLREQFGNK-TIAPNOSGGDPPIVMHSE 354
DB 329 TTGEIIGDIRQAHCTLRNTEWNTLAKITEKLEQFGNNTIIVFNHSSGGDPPIVMHSE 388
QY 355 CGGEFFCYCNSQTFNSAWNTSGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAY 414
DB 389 CGGEFFCYCNSQTFNSAWNTSGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAY 444
QY 415 ALPIKILIRCSNITGLLITRDGGG--ENQTTIFRPGGDMRDNRSELYKYKVKVIEP 472
DB 445 APPIRQIRCSNITGLLITRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 504
QY 473 LGVAPTKARRVVQREKRAVGMFLGAGSTMGATSMALTVQARQLLSGIVQOQ 532
DB 505 LGVAPTKARRVVQREKRAVGMFLGAGSTMGATSMALTVQARQLLSGIVQOQ 564
QY 533 NLLRAIAQAHLLQLTVMGILQARILAVERYLKDQQLLGFPGWCGKLICTTAVPWA 592
DB 565 NLLRAIAQAHLLQLTVMGILQARILAVERYLKDQQLLGFPGWCGKLICTTAVPWA 624
QY 593 SWSNKTLDQIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLQDKWASL 649
DB 625 SWSNKSLEIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLQDKWASL 681

RESULT 11
Q9WJU4 PRELIMINARY; PRT; 863 AA.
ID Q9WJU4
AC Q9WJU4
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
EN Envelope polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,
RA Burger H.,
RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to
RT Non-Progressive Infection."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69592; AAD10939.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane.
SQ SEQUENCE 863 AA; 98269 MW; BE12C858E9CE4D11 CRC64;

Query Match 82.3%; Score 2858.5; DB 15; Length 863;

Best Local Similarity 81.7%; Pred. No. 1.5e-226;
Matches 537; Conservative 44; Mismatches 53; Indels 23; Gaps 5;

QY 3 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62
DB 33 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92
QY 63 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 115
DB 93 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 149
QY 116 IIVVMEQKGMKRNCSFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNTKYL 173
DB 150 -WGMERGEIKNGSFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNTKYL 205
QY 174 NTSVITQACPKVSFEPIPIHYCTPTGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 233
DB 206 NTSVITQACPKVSFEPIPIHYCAPAGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 265
QY 234 VSTQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTNRKRVTLGPGRV 293
DB 266 VSTQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTNRKRVTLGPGRV 325
QY 294 WYTGELILGNTRQAHNCNISRAQWNTLQOIATTLREQFG-NKTIAENQSGGDPPIVMH 352
DB 326 WYTGELILGNTRQAHNCNISRAQWNTLQOIATTLREQFG-NKTIAENQSGGDPPIVMH 385
QY 353 FNCGGEFFCYCNSQTFNSAWNTSGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGK 412
DB 386 FNCGGEFFCYCNSQTFNSAWNTSGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGK 439
QY 413 MYALPIKILIRCSNITGLLITRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 472
DB 440 MYALPIKILIRCSNITGLLITRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 499
QY 473 LGVAPTKARRVVQREKRAVGMFLGAGSTMGATSMALTVQARQLLSGIVQOQ 532
DB 500 LGVAPTKARRVVQREKRAVGMFLGAGSTMGATSMALTVQARQLLSGIVQOQ 559
QY 533 NLLRAIAQAHLLQLTVMGILQARILAVERYLKDQQLLGFPGWCGKLICTTAVPWA 592
DB 560 NLLRAIAQAHLLQLTVMGILQARILAVERYLKDQQLLGFPGWCGKLICTTAVPWA 619
QY 593 SWSNKTLDQIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLQDKWASL 649
DB 620 SWSNKSLEIWNMTWMEWDREIDNTHLYTLIEESQOQKNOQELLQDKWASL 676

RESULT 12
Q9DL81 PRELIMINARY; PRT; 858 AA.
ID Q9DL81
AC Q9DL81
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,
RA Greenberg M.L.;
RT "Evolution of the Human Immunodeficiency Virus Type 1 Envelope during
RT Infection Reveals Molecular Correlations of Specificity for Coreceptor
RL Utilization and AIDS Pathogenesis."
RL J. Virol. 74:11858-11872 (2000).
DR EMBL; AF310126; AAG37004.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 858 AA; 97823 MW; CC028DC1178368B CRC64;

Query Match 82.2%; Score 2857; DB 15; Length 858;
Best Local Similarity 82.5%; Pred. No. 2e-226;
Matches 536; Conservative 44; Mismatches 56; Indels 14; Gaps 7;

QY 3 LWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTENF 62
DB 33 LWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTENF 92

QY 63 NMWKNMVQEHEDIIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTITLSIIVWEQ 122
DB 93 NMWKNMVQEHEDIIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTITLSIIVWEQ 148

QY 123 RGKEMRNCSENIITTSIRDKVQREYALFKDVEPIDDKNTVNTK---YRLINCNTSV 179
DB 149 EKGKIKNCSENIITGIRGVQ-EYSLFKLDVPIIDSRNNSNSTEYNSYRLINCNTSV 207

QY 180 ITQACPVSFEPIPIHYCTPTGTFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQL 239
DB 208 ITQACPVSFEPIPIHYCAPAGFAILKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQL 267

QY 240 LNSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGE 299
DB 268 LNSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGE 326

QY 300 ILGRIQAHNISRAQWNTLQOIIATLREDFGNTKIAPNOSGGDPPIVMSFCGGEF 359
DB 327 ILGRIQAHNISRAQWNTLQOIIATLREDFGNTKIAPNOSGGDPPIVMSFCGGEF 386

QY 360 PYCNSTQIFNSAWNTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGMKAMVALPIK 419
DB 387 PYCNSTQIFNSAWNTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGMKAMVALPIK 441

QY 420 GLIRCSSNITGLLLTRDGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 479
DB 442 GLIRCSSNITGLLLTRDGGNGQTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 501

QY 480 AKRRVQREKRAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 539
DB 502 AKRRVQREKRAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 561

QY 540 KAOQHLLQLTWGIKQOLQARILAVRYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 599
DB 562 KAOQHLLQLTWGIKQOLQARILAVRYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 621

QY 600 DOIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 649
DB 622 DOIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 671

RESULT 13
Q9J022 PRELIMINARY; PRT; 853 AA.
AC Q9J022;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BORI;
RA MEDLINE=20091324; PubMed=10623731;
RA Shieh J.T., Martin J., Baltuch G., Malim M.H., Gonzalez-Scarano F.;
RT "Determinants of syncytium formation in microglia by human
RT immunodeficiency virus type 1: role of the V1/V2 domains.";

J. Virol. 74:693-701 (2000).
DR EMBL; AF227273; AAF61104.1; -
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 853 AA; 96346 MW; 1601A1A72468E377 CRC64;

Query Match 82.2%; Score 2855.5; DB 15; Length 853;
Best Local Similarity 82.8%; Pred. No. 2.7e-226;
Matches 538; Conservative 41; Mismatches 54; Indels 17; Gaps 7;

QY 2 NLWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTENF 61
DB 32 NLWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTENF 91

QY 62 FMWKNMVQEHEDIIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTITLSIIVWE 121
DB 92 FMWKNMVQEHEDIIISLDQSLKPCVKLTPLCVTLNCTDY-LNNAINTTSSK-----E 145

QY 122 QRGKEMRNCSENIITTSIRDKVQREYALFYKLDVEPIDDKNTTNTTYRLINCNTSVIT 181
DB 146 GEMGEIKKCSFNVTIRDKVQREYALFYKLDVVPVIGN-----GNTSYRLINCNTSVIT 200

QY 182 QACPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLN 241
DB 201 QACPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLN 260

QY 242 GSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGEIL 301
DB 261 GSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKSIHIGPGRFTYTTGDI 320

QY 302 GNIQAHNISRAQWNTLQOIIATLREDFG-NKTIAPNOSGGDPPIVMSFCGGEF 360
DB 321 GDIQAYCILSRQWNTLQOIIATLREDFG-NKTIAPNOSGGDPPIVMSFCGGEF 380

QY 361 YCNSTQIFNSAWNTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGMKAMVALPIK 420
DB 381 YCNSTQIFNSAWNTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGMKAMVALPIK 436

QY 421 LIRCSSNITGLLLTRDGGNGQ-TTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 479
DB 437 QIKCSSNITGLLLTRDGGNGSGITTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 496

QY 480 AKRRVQREKRAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 539
DB 497 AKRRVQREKRAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 556

QY 540 KAOQHLLQLTWGIKQOLQARILAVRYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 599
DB 557 KAOQHLLQLTWGIKQOLQARILAVRYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 616

QY 600 DOIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 649
DB 617 DOIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 666

RESULT 14
Q9E1R7 PRELIMINARY; PRT; 855 AA.
AC Q9E1R7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Envelope protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20442410; PubMed=10984542;

RA Cleghorn F.R., Jack N., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDanal C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RL Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF277071; AAG22514.1; -;
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 855 AA; 97102 MW; 26271D6CCCCCFAC CRC64;

Query Match 82.2%; Score 2855.5; DB 15; Length 855;
Best Local Similarity 82.0%; Pred. No. 2.7e-226;
Matches 534; Conservative 42; Mismatches 52; Indels 23; Gaps 6;

QY 4 WTVTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 63
Db WTVTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 95
QY 64 MKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNC-----TDLNTNTNTTSLIIV 118
Db MKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNC-----TDLNTNTNTTSLIIV 118
QY 96 MKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 155
Db MKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 155
QY 119 VMEQKGMKRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTS 178
Db VMEQKGMKRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTS 178
QY 156 T-----GEMKNCSPNITTSVRDKVREYALFYKLDVTPIDDN---NDNSGYLLNCNTS 206
Db T-----GEMKNCSPNITTSVRDKVREYALFYKLDVTPIDDN---NDNSGYLLNCNTS 206
QY 179 VITQCPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 238
Db VITQCPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 238
QY 207 VITQCPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 266
Db VITQCPKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 266
QY 239 LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 298
Db LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 298
QY 267 LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
Db LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
QY 299 EILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 358
Db EILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 358
QY 327 EILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 386
Db EILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 386
QY 359 FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 418
Db FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 418
QY 387 FFYCNSTQLFNSTWNT--NDT-----RNDTGNITLPCRIRQIVNRQEVGKAMYAPPI 438
Db FFYCNSTQLFNSTWNT--NDT-----RNDTGNITLPCRIRQIVNRQEVGKAMYAPPI 438
QY 419 KGLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 478
Db KGLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 478
QY 439 RQGISCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 498
Db RQGISCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 498
QY 479 KARRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 538
Db KARRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 538
QY 499 KARRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 557
Db KARRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 557
QY 539 IKAQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 598
Db IKAQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 598
QY 558 IEAQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKS 617
Db IEAQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKS 617
QY 599 LDOIWNMTWMDREIDNTYLIYTLIBESQOQKQELLDKWSL 649
Db LDOIWNMTWMDREIDNTYLIYTLIBESQOQKQELLDKWSL 649
QY 618 MDDIWNMTWMDREIDNTYLIYTLIBESQOQKQELLDKWSL 658
Db MDDIWNMTWMDREIDNTYLIYTLIBESQOQKQELLDKWSL 658

RESULT 15
O9DH32 PRELIMINARY; PRT; 858 AA.
ID O9DH32
AC Q9DH32;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Envelope protein).
GN ENV.
OS Human immunodeficiency virus 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QH1420C.2;
RX MEDLINE=20541986; PubMed=11090186;
RA Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,
RA Greenberg M.L.;
RA "Evolution of the Human Immunodeficiency Virus Type 1 Envelope during
RT Infection Reveals Molecular Corollaries of Specificity for Coreceptor
RT Utilization and AIDS Pathogenesis.";
RL J. Virol. 74:11858-11872(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Jack N., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDanal C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RL Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF310124; AAG37002.1; -;
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 858 AA; 97843 MW; 6B77091DC964BF94 CRC64;

Query Match 82.2%; Score 2854; DB 15; Length 858;
Best Local Similarity 82.6%; Pred. No. 3.6e-226;
Matches 537; Conservative 42; Mismatches 57; Indels 14; Gaps 7;

QY 3 LMWTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 62
Db LMWTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 92
QY 63 NWKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 122
Db NWKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 122
QY 93 NWKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 148
Db NWKNNVVEQMHEDIISLWQSLKPCVKLPLCVTLNCIDYVKNNDTNNSTNNSSVME 148
QY 123 RKGGMKRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTSV 179
Db RKGGMKRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTSV 179
QY 149 EOKGEIKNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTSV 207
Db EOKGEIKNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNKKYLLNCNTSV 207
QY 180 ITQACPCKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 239
Db ITQACPCKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 239
QY 208 ITQACPCKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 267
Db ITQACPCKVSFEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 267
QY 240 LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 299
Db LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 299
QY 268 LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
Db LINGSAAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
QY 300 ILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 359
Db ILGNIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 359
QY 327 IIGDIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 386
Db IIGDIRQAHNCISRAQWNTLQOIATTLREQFNKTIAPNOSGGDPEIWMHSPNCGGE 386
QY 360 FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 419
Db FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 419
QY 387 FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 441
Db FYCNSITQLFNSAWNTSNGTSVTRKQDGTIIILPCRIKQIINRWQVVKAMVALPI 441
QY 420 GLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 479
Db GLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 479
QY 442 GLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 501
Db GLIRCSSNITGLLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIEPLGVAPT 501
QY 480 AKRRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 539
Db AKRRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 539
QY 502 AKRRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 561
Db AKRRVYVQREKRAVGMGLGFLGAAGSTMATVQARQLLSGIVQOQNLLRA 561
QY 540 KAOQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 599
Db KAOQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 599
QY 562 EAOQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 621
Db EAOQHLQLTVMGIIKQIARILAVERYLKDQQLLGFPGWCGSKLICCTAVPWNASNKT 621

Qy 600 DQIWNNTWMEWDREIDNVTHTLYTLIESONQOEKNOELLOLDKWASL 649
Db 622 TDWNTWMEWDREITNTYELIYTLLEDAQNQOEKNEODLLELDKWASL 671

Search completed: January 13, 2004, 09:11:26
Job time : 44 secs